University of Southampton

Faculty of Engineering and Physical Science

Civil, Maritime, and Environmental Engineering Department

The Driver State Monitoring in Semi-Automated Vehicles- the Influence of the Circadian Phase

by

Sylwia Izabela Kaduk

ORCID ID https://orcid.org/0000-0003-0730-2389

Thesis for the degree of Doctor of Engineering

[December 2020]

University of Southampton

Abstract

Faculty of Engineering and Physical Science

Civil, Maritime, and Environmental Engineering Department

Thesis for the degree of Doctor of Engineering

The Driver State Monitoring in Semi-Automated Vehicles- the Influence of the Circadian Phase

by

Sylwia Izabela Kaduk

Background: automation offers the potential to mitigate or reduce the risks related to driving. However, there are also some new challenges for drivers related to semi-automated driving. Two phases of semi-automated driving that raised concerns of the researchers were a period of automation that requires a monitoring activity from the driver and the take-over of manual control following the automated mode. Topic: the aim of this doctoral thesis was to propose models of the driver state monitoring in semi-automated vehicles and present data on the psychophysiological changes occurring during semi-automated driving, as well as the circadian effect on semi-automated driving and driver state monitoring. Methods: fifty-two participants were recruited to the experiment on semi-automated driving. They participated in two experimental sessions day-time session (9 a.m.- 1 p.m.) and a night-time session (10 p.m.- 2 a.m.). They went through the experimental scenario simulating semi-automated driving with phases of manual driving, automated phase, take-over and manual driving. During the experiment their psychophysiological functions were recorded with the following measures: electrooculography, electromyography, electrocardiography, respiration belt, electrodermal activity device, oximetry for the pulse and blood oxygenation, their voice was recorded for the acoustic voice analysis, saliva was collected for the hormonal analysis, and four questionnaires were collected at different stages of the experiment. Additionally, electroencephalography was recorded; however, its analysis was not included in this thesis. Results: two predictive models were proposed to predict performance after take-over and attention during automation. Analysis of the time-course of the semi-automated driving suggested a decrease of the driving performance after automation associated with increased sleepiness, increased fatigue, decreased readiness to take-over and decreased mental workload. Some physiological changes suggested mental underload. Comparison of the circadian phases resulted in multiple physiological, behavioural and cognitive changes. Conclusions: physiology can be used to predict the driver's performance in semi-automated vehicles; however, the proposed models are not ready to be implemented in the cars. Automation creates a risk for driving safety due to mental underload. Sleepiness and fatigue present the largest risk for automation monitoring, while suboptimal mental workload and arousal for the safety of the take-over. The circadian phase affects the psychophysiology and performance of the driver; however, the direction of the effects requires further investigation.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

RESEARCH THESIS: DECLARATION OF AUTHORSHIP

Print name: Sylwia Izabela Kaduk

Title of thesis: The Driver State Monitoring in Semi-Automated Vehicles- the Influence of the Circadian Phase

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

   Journal Publications:

Sylwia I. Kaduk, Aaron P. J. Roberts & Neville A. Stanton (2020) Driving performance, sleepiness, fatigue, and mental workload throughout the time course of semi-automated driving—Experimental data from the driving simulator. Human Factors and Ergonomics in Manufacturing & Service Industries, DOI: 10.1002/hfm.20875

Sylwia I. Kaduk, Aaron P. J. Roberts & Neville A. Stanton (2020) The circadian effect on psychophysiological driver state monitoring, Theoretical Issues in Ergonomics Science, DOI: 10.1080/1463922X.2020.1842548

Sylwia I. Kaduk, Aaron P. J. Roberts & Neville A. Stanton (2020) The manual shift in phase: the impact of circadian phase on semi-autonomous driving. What can we learn from current understanding in manual driving?, Theoretical Issues in Ergonomics Science, DOI: 10.1080/1463922X.2020.1758829

Conference Publications:

Kaduk S. I., Roberts A. P., J., Stanton N. A., (2020) I just woke up behind the wheel- Sleep Inertia as a new risk in semi-automated cars. In: Contemporary Ergonomics and Human Factors 2020. Eds. Rebecca Charles and Dave Golightly. CIEHF.

Kaduk S.I., Roberts A.P., Stanton N.A. (2021) Psychophysiological Predictors of Motion Sickness in the Driving Simulator. In: Cassenti D., Scataglini S., Rajulu S., Wright J. (eds) Advances in Simulation and Digital Human Modeling. AHFE 2020. Advances in Intelligent Systems and Computing, vol 1206. Springer, Cham

Kaduk S., Roberts A., Stanton N. (2020) Surviving in Frankenstein's Lab- How to Build and Use a Laboratory that Applies Multiple Physiological Sensors. In: Cassenti D. (eds) Advances in Human Factors and Simulation. AHFE 2019. Advances in Intelligent Systems and Computing, vol 958. Springer, Cham.

Signature: Date:01.12.2020

ACKNOWLEDGEMENTS

For more than three years, I have been dedicating the majority of my time, thoughts, and effort to this research project. It was a unique opportunity and I am grateful that I was able to do that. I feel that my mind developed during that period in ways and directions that I could never imagine. I feel richer and more fulfilled as a person. Hopefully, I will also be able to offer the skills and knowledge that I have gained to the world and make it a better place. Many people contributed to this opportunity and the successful end of this work. I am grateful to all of them. I would like to mention some of them to underline their contribution.

First, I would like to thank my husband, Tobiasz Kaduk. He was the one who convinced me to become a scientist. He travelled the world with me, inspired me, and supported me on every stage of this journey. He should be a co-author of this dissertation as he contributed so much with his love, enthusiasm and wisdom.

My amazing parents are both owed a great thank you.

My mom, Agnieszka Wojas, was supportive in many ways and she always repeated that I am going to make it, even when I did not believe it.

My father, Włodzimierz Wojas, shared his knowledge in mathematics, data visualisation and graph theory, while patiently discussing my dissertation with me again and again.

I am grateful to scientists from the Surrey Sleep Research Centre, who taught me a lot and included me in their fascinating research when I was still an inexperienced beginner. Professor Dirk-Jan Dijk, Dr Nayantara Santhi, and Dr Maria-Angeles Bonmati-Carrion all significantly contributed to my understanding of sleep and circadian rhythmicity. It is thanks to them that I have studied the circadian influence on semi-automated driving.

I am also grateful to my yoga teachers Rochelle Parkes and Gemma Powell, whose never-ending enthusiasm and yoga insight showed me how to sustain a balanced state of mind and allowed me to develop spiritually during the last years.

Finally, I want to thank my supervisors, professor Neville Stanton and Dr Aaron Roberts, the whole Human Factors Engineering team and the Transportation Research Group at the University of Southampton for creating an environment where I was able to thrive. Special thanks are owed to Dr Jediah Clark and James Brown. They made this experience so positive for me by many funny and inspiring conversations, all the support in technical and computer-related problems, and just being around with their positive and open attitude.

DEFINITIONS AND ABBREVIATIONS

ANN- artificial neural network

ECG- electrocardiography

EDA- electrodermal activity

EEG- electroencephalography

EMG- electromyography

EOG- electrooculography

ERP- event-related potential

fNIRS- functional near-infrared spectroscopy

HI:DAVe- Human Interaction: Designing Autonomy in Vehicles

HRV- heart rate variability

KSS- Karolinska sleepiness scale

NASA-TLX- NASA task load index

NN50%- number of pairs of successive NN (R-R) intervals that differ by more than 50 ms

PERCLOS- Percentage closed

PNN50%- the proportion of NN50 divided by the total number of NN (R-R) intervals

RMSSD- root mean square of successive differences between normal heartbeats

RSA- respiratory sinus arrhythmia

SCL- skin conductance level

SDSD- standard deviation of successive RR interval differences

SNR- signal to noise ratio

SVM- support vector machine

TASCC- Towards Autonomy – Smart and Connected Control

TORS- take-over readiness scale

XLR- external line return

1. INTRODUCTION

1.1 BACKGROUND

This doctoral research formed part of a larger project called HI:DAVe (Human Interaction: Designing Autonomy in Vehicles). HI:DAVe was a collaboration between the University of Southampton, Cambridge University and Jaguar Land Rover with funding from the Engineering and Physical Sciences Research Council. The project examined the use of semi-automated vehicles from a Human Factors perspective. The critical topic of the investigation was the transition between semi-automated driving and manual driving, the so-called take-over problem.

The take-over is perceived as one of the potential risk sources in semi-automated driving, and therefore many investigators put effort into unravelling its dynamics and characteristics. It motivated the joint effort to understand semi-automated driving take-over from a human factors perspective.

A wide variety of literature suggested that semi-automated driving can alter the state of the driver (Kyriakidis et al., 2019), potentially impacting their capacity to safely resume manual control of the vehicle. Partial automation shifts the role of the driver from active vehicle control to that of passive monitoring activities. Unfortunately, monitoring is a role challenging from the perspective of human cognitive systems (Kyriakidis et al., 2019). It requires sustained attention but does not provide sufficient mental stimulation (Kyriakidis et al., 2019; Oken et al., 2006; Warm et al., 2008). Therefore, it might lead to a reduction in situation awareness (‘falling out of the control loop’), ‘mind-wandering’, fatigue or even falling asleep (Stanton, 2015). People might also choose to engage in a more stimulating task and neglect driving safety monitoring (Dogan et al., 2017; Eriksson & Stanton, 2017; Kyriakidis et al., 2019). Several states, including distraction, lack of situation awareness, or fatigue have been demonstrated the negative impact of semi-automated vehicles on driving behaviour (Heikoop et al., 2016; Parnell et al., 2016). Moreover, people tend to wrongly assess their fitness to drive (Filtness et al., 2017; Ftouni et al., 2013; Howard et al., 2014). Therefore, it is necessary to understand the driver state before take-over, as well as develop methods of assuring the safety of the process.

One possible method to alleviate the risk identified above could be using information derived from driver psychophysiology to detect the driver state and readiness to take over manual control of the vehicle. However, the ability to effectively measure the state of the driver in relation to driving performance is a relatively neglected area of research. Also, information about the application of the driver state monitoring systems in semi-automated vehicles are scarce.

The main aim of this research work was to evaluate a wide range of methods for psychophysiological driver state monitoring in the semi-automated vehicle. The evaluation was meant to give a possibility to compare the effectivity of different methods as well as to assess their potential value in the prediction of actual driving performance. This main aim was accompanied by the analysis of the patterns in driver state during semi-automated driving and circadian effect on driver state monitoring and driver performance in semi-automated vehicles.

1.2 RESEARCH MOTIVATION

The primary motivation of this research was to increase road safety in the perspective of the newly introduced technology of semi-automated vehicles. New technologies, including artificial intelligence, bring exciting opportunities to the society. It is the responsibility of scientists and manufacturers to ensure their safety and sustainability. This work experimentally investigated the physiology, driving performance, and psychological state of the driver in a simulated semi-automated vehicle. The main goal of this investigation was to evaluate and compare the practical value of different methods of driver state monitoring in the semi-automated vehicle but it also offered an analysis of the circadian effect on semi-automated driving, the circadian effect on driver state monitoring and analysis of general patterns of driver state in the semi-automated vehicle. It has been proposed that approaches to driver state monitoring should be investigated to provide a better understanding of the psychophysiological processes together with the effects of the circadian rhythm on semi-automated driving.

The topic of semi-automated driving is relatively new and the knowledge about driver state in the process of automation and take-over is mainly based on predictions, rather than empirical science (Kyriakidis et al., 2019). There were many attempts to assess different physiological recordings methods in the context of driver state monitoring or risky state detection; however, they were mostly evaluated in the context of manual driving. They allowed associating many physiological states with driving risks. For example, the analysis of spectral features in ECG and EMG with KNN classifier allowed to obtain even 96.75% and 92.31 accuracy in drowsiness and distraction detection (Sahayadhas et al., 2015). Electrodermal liability in sleep-deprived people allowed to predict drowsiness (Michael et al., 2012), as well as neural network-based LDS analysis of EOG signal (Zhu et al., 2014). However, the majority of the studies lacked an ecological validity for the driving environment. To validate the accuracy of their predictions, they used other physiological measures, questionnaires, experimental manipulation in the non-driving laboratory or basic cognitive tasks not related to driving. For example, Sahayadhas et al. (2015) created a number of experimental conditions to induce distraction and sleepiness. Behavioural distraction was induced with reading and answering the text message while driving the driving simulator, cognitive distraction by talking on the phone and solving mathematical equations, while drowsiness by a long period of the experiment. The high accuracy of classification that was reported was classification between different stages of the experiment. These results showed that ECG and EMG significantly differed between the situations when people answered the text message or talked on the phone. It brought an important knowledge about physiology but did not imply that distraction can be detected that way in a driver and moreover that it can be a predictor of driver performance. Another example is an experiment by Michael et al. (2012). They validated the level of drowsiness in sleep-deprived participants with a subjective scale and observation of sleepiness symptoms. Their results indicated electrodermal liability as a strong correlated of subjective sleepiness of participants, which still does not immediately apply to the driver performance. Zhu et al. (2014) used a simple button-press task and managed to predict results with EOG data, which does not mean that this could be immediately applied in the prediction of the performance in a complex activity that is driving. Such methods of validation confirmed the possibility of state detection with psychophysiology but there is a need for additional research investigating its ecological validity in the driver state monitoring. One of the reasons is that driving performance does not always change with the state of the driver. For example, a sleepy or distracted driver might still drive well using their additional available cognitive capacity (Parasuraman et al., 2008; Ross et al., 2014; Young & Stanton, 2002). The current study aimed to predict performance and bring more ecological value that could be applied in real-life situations. It also used a semi-automated driving experimental scenario to evaluate the methods in this type of vehicle.

Psychophysiological monitoring is a complicated area because it requires a detailed collection of physiological data in an environment that creates a lot of signal noise. A proper selection of factors in complex physiological signals is needed, as well as for analytic methods that could derive meaningful information from the complex signals. In the case of this work, over ten psychophysiological functions were recorded simultaneously, which created an extremely difficult experimental environment as well as a challenging analytic task due to the need for features selection. This added more experimental data to support or confront the theoretical predictions of the effects of automation.

1.3 RESEARCH OUTCOMES AND HYPOTHESES

The purpose and the main topic of this research thesis was driver state monitoring in semi-automated driving. The experiment was designed to create an opportunity to compare the effectivity of multiple physiological methods of driver state monitoring and to evaluate their potential in the prediction of actual driver performance, rather than the subjective state. The psychophysiological experiment in the driving simulator set-up offered the opportunity to investigate psychophysiological changes during the time-course of semi-automated driving as well as the circadian effect on the driver state monitoring. This project investigated four main research questions:

1. What are the physiological measures and strategies of analysis that could allow effective monitoring of the driver state before manual driving take-over and prediction of the driving performance after the take-over?
2. What are the physiological measures and strategies of analysis that could allow effective monitoring of the driver state during the automated phase to assess if the driver is paying appropriate attention?
3. What is the general tendency of the psychophysiological driver state during the time-course of semi-automated driving (manual driving, automated phase, take-over, and manual driving)?
4. What is the effect of the circadian rhythmicity on the driver state monitoring during semi-automated driving?

The primary aim of this research was the identification of an effective and accurate system of psychophysiological monitoring of the driver in semi-automated vehicles. Semi-automated vehicles are still a relatively new technology and therefore understanding the mechanisms that could assure safety would be both crucial and novel. Driving safety, no matter if manual or semi-automated, is a topic of global importance. World Health Organization identified vehicle collisions as an epidemic and united with ministers in over one-hundred countries. Ministers agreed to halve road-related deaths by 2030 in the 3rd Global Conference on the Road Safety (*Ministers to Agree New Global Road Safety Agenda to 2030*, n.d.). Automated driving could help in this endeavour.

The circadian perspective in this research was motivated by two lines of thought. Night-time manual driving was documented as being significantly riskier; however, the circadian effect on driving has been generally neglected despite evidence (Akerstedt et al., 2013; Matthews, Ferguson, Zhou, Sargent, et al., 2012; Mitler et al., 1988). A similar phenomenon could be observed in semi-automated driving and the topic is worthy of investigation. The second reason is that circadian rhythmicity influences many physiological functions (Blatter & Cajochen, 2007; Dijk et al., 1992; Van Dongen & Dinges, 2000) and as so, could also affect the interpretation of the signals in the system of physiological monitoring of the driver.

The research aimed to offer both innovation for semi-automated vehicles and understand the processes related to semi-automated driving from the perspectives of physiology and circadian rhythms.

1.3.1 RESEARCH HYPOTHESES

The following research hypotheses were based on the literature reviews:

1. Psychophysiological measurements of the driver during semi-automated driving can provide a prediction of the driving performance after take-over.
    a. Sub-hypothesis: psychophysiological indicators that can predict driving performance are related to one of the risky states identified in chapter 2, namely sleepiness, fatigue, distraction, mental workload or situation awareness.
2. Psychophysiological measurements of the driver during semi-automated driving can provide a prediction of their attention during the automated mode of semi-automated driving.
    a. Sub-hypothesis: psychophysiological indicators that can predict driver attention are related to one of the risky states identified in chapter 2, namely sleepiness, fatigue, distraction, mental workload or situation awareness.
3. Driver psychophysiological state and performance differ before and after automated mode.
4. Driver performance is worse after automated mode, while their psychophysiological state is related to the lower cognitive state.
5. Driver psychophysiological state and performance in semi-automated vehicles differ between day and night.
    a. Sub-hypothesis: driving performance and attention during automation decrease during the night.

1.3.2 RESEARCH OUTCOMES

Modelling of psychophysiological and performance data led to predictive models that partially explain changes in driver performance and attention during semi-automated driving. The models did not explain enough variance to be implemented in the vehicles at the present moment; however, they provided a base for future research and knowledge about risky driver states for different stages of semi-automated driving.

The comparison of the driver performance and the driver state during the time-course of semi-automated driving suggested that driving performance was worse after automation due to cognitive underload.

The comparison of the driver performance and the driver state between day and night experimental sessions allowed to identify several circadian differences; however, the assumed direction of worse performance at night was not confirmed.

1.4 THESIS STRUCTURE

This thesis contains eleven chapters and nine appendices. Chapter one is the introductory chapter and chapters two to six present findings from various reviews of the literature. The literature reviews provide an overview of the research landscape as well as identify gaps in knowledge. The reviews have also helped to identify driver states that jeopardize driving safety, available methods of driver state monitoring, and the existing information about the circadian effect on manual driving, semi-automated driving, as well as driver state monitoring. As the primary strategy of all the five literature reviews has been the same, it was described once below to avoid repetitions. Only the details, such as databases or inclusion/exclusion criteria, were defined separately for each case.

The methodology of the literature reviews was based on the grounded theory approach, which was previously used in the driving research (Parnell et al., 2016), as well as the other human factors areas (Rafferty et al., 2010). Grounded theory was used in exploratory social research when the hypotheses were developed in the process of the literature search. The literature was compared continuously to establish search directions. The process was based on induction, deduction, and verification. The researcher was looking for questions and patterns emerging from the literature. The method used a strategy of open, axial and selective coding to identify main themes or factors for a given problem (Heath & Cowley, 2004). Such an approach was beneficial in the case of this work because it allowed a broader view not overly restrained by the presumptions.

Chapter seven presented a more technical and practical description of the laboratory construction and set of recommendations for the simultaneous measurements of multiple psychophysiological functions. Psychophysiological experiments have a variety of requirements to ensure the high quality of the data (Cacioppo et al., 2007; Cutmore & James, 1999). There was a need for a laboratory construction to minimize acoustic noise, electromagnetic noise and to provide a suitable environment for other measures, for example, a freezer for saliva or a driving simulator. The noise insulated Faraday Cage with a low fidelity driving simulator was built in the Transportation Research Group garage.

One elaborate experiment was conducted during this doctorate and different perspectives of the analysis led to the results presented in the chapters from eight to ten. Chapter eleven gathered results shown in all the chapters, discussed them and gave recommendations based on the experimental and theoretical work, as well as future research directions.

1.4.1 CHAPTER 2: RISKY DRIVER STATES IN SEMI-AUTOMATED VEHICLES- REVIEW

Chapter 2 reviewed literature related to the risky driver states to identify the states that could be interesting for a driver state monitoring system. The most commonly described states were drowsiness, fatigue, behavioural distraction, and cognitive distraction (Caird, 2015; Jackson, Raj, et al., 2016; Johns, 2000; Liang & Lee, 2010); however, most of the literature was related to the manual driving. Based on the Consensus Model (Heikoop et al., 2016) also suboptimal mental workload and insufficient situation awareness were treated as relevant risks in semi-automated driving. The chapter also described risks coming from sleep inertia (Ferrara & De Gennaro, 2000; Wörle et al., 2020) that was identified as a new driving risk emerging from vehicle automation.

1.4.2 CHAPTER 3: METHODS OF THE DRIVER STATE MONITORING-REVIEW

Chapter 3 presented a review of the literature related to driver state monitoring or risky state detection. The aim was to identify available psychophysiological measures and select the most promising for the experimental work. An abundant list of methods used for the driver state monitoring was identified: electroencephalography (EEG) (Dhupati et al., 2010), Hybrids of Methods (Sahayadhas et al., 2013), Eye-Tracking (Hogervorst, Brouwer, & van Erp, 2014), electrocardiography (ECG) (Brookhuis et al., 1991), electrooculography (EOG) (Borghini et al., 2014), functional near-infrared spectroscopy (fNIRS) (Aranyi et al., 2015), electrodermal activity (EDA) (Miyake et al., 2009), Acoustic Voice Analysis (Krajewski, Batliner, et al., 2009), event-related potential (ERP) (Resalat et al., 2012), electromyography (EMG) (Oken et al., 2006), Questionnaires (Horne & Baulk, 2004), Blood Pressure (Veltman & Gaillard, 1996), Infrared Video Camera (Vitabile et al., 2010), Facial Expression (Fan et al., 2010), Saliva Analysis (Zeier et al., 1996), Body Temperature (Milosevic, 1997), Pupillometry (Mitler et al., 1988), Respiration (Rodríguez-Ibáñez et al., 2011), Driving Performance (Bando & Nozawa, 2015), Body Position (Van Dongen & Dinges, 2000), Head Movements (Murata et al., 2015), Oximetry (Sharma & Bundele, 2015), Actigraphy (Mullaney et al., 1980), Blood Glucose (Fairclough & Houston, 2004), and Doppler Flow Meter (Miyake et al., 2009). The most commonly used measures were discussed in detail.

1.4.3 CHAPTER 4: CIRCADIAN EFFECT ON MANUAL DRIVING- REVIEW

Chapter 4 reviewed existing knowledge about the influence of the circadian phase on manual driving. Night and afternoon were times of documented decrements in driving performance (Akerstedt et al., 2013; Lowden et al., 2009); however, the methodology used in the majority of the studies did not allow the identification of circadian phase as the sole cause, as it is challenging to dissociate it from sleep deprivation (Blatter & Cajochen, 2007).

1.4.4 CHAPTER 5: CIRCADIAN EFFECT ON SEMI-AUTOMATED DRIVING- REVIEW

It was documented that manual driving was significantly riskier at night. However, there are no experimental data about the influence of the circadian rhythm on semi-automated driving. Chapter 5 used a consensus model of the driver in automation (Heikoop et al., 2016) and provided evidence that multiple factors in the model could be affected by circadian rhythm. A multi-period consensus model of the driver in automation was proposed (Kaduk et al., 2020) to create a theoretical basis for experimental research about the circadian effect on semi-automated driving.

1.4.5 CHAPTER 6: CIRCADIAN EFFECT ON DRIVER STATE MONITORING- REVIEW

Systems of the driver state monitoring use various psychological observations or physiological signals to derive information about the driver state and predict risks (Melnicuk et al., 2016). Many physiological and psychological states undergo circadian rhythmicity (Blatter & Cajochen, 2007; Dijk et al., 1997), and therefore have a different baseline at different circadian phases. Chapter 6 used a literature review to support the suggestion that interpretation of the information from the driver state monitoring should take a circadian phase into account.

1.4.6 CHAPTER 7: CREATION OF THE LABORATORY FOR PSYCHOPHYSIOLOGICAL MEASUREMENTS OF THE DRIVER

The experimental data gathered during this doctorate presented a tremendous technical challenge. Multiple physiological functions were recorded simultaneously during the simulated semi-automated drive. For this purpose, a laboratory was constructed in the garage of Transportation Research Group. A driving simulator was placed inside a noise insulated Faraday Cage. Several steps were undertaken to assure the best quality of data recording and reduce the amount of noise in signals. Chapter 7 described the laboratory construction and provided recommendations for a laboratory measuring multiple physiological functions simultaneously. A decision tree was presented as a prototype of a research support tool with a unique set of recommendations for laboratory and experimental set-up for different combinations of physiological recordings used in the experiment.

1.4.7 CHAPTER 8: TIME-COURSE OF SEMI-AUTOMATED DRIVING- EXPERIMENTAL RESULTS

Chapter 8 provided details of the experimental work conducted during this doctorate and analysis related to the time-course of semi-automated driving. The results supported the hypotheses suggesting that driving performance decreased after automation, participants felt sleepier, more fatigued, and less ready to take-over the manual control over the vehicle. Their physiology suggested a cognitive underload that could explain the decrease in their performance. Moreover, participants were not able to accurately assess their own fitness to drive.

1.4.8 CHAPTER 9: CIRCADIAN EFFECT ON SEMI-AUTOMATED DRIVING AND DRIVER STATE MONITORING- EXPERIMENTAL RESULTS

Three literature reviews suggested a potential decrease in driver performance in semi-automated vehicles at night. Chapter 9 provided results of the statistical analysis of the experimental data, comparing day and night driving performance and psychophysiological states. The results only partially supported the hypotheses. There were multiple circadian differences that could affect the driver state monitoring system; however, the direction of the changes was not clear, and the topic requires further research.

1.4.9 CHAPTER 10: DRIVER STATE MONITORING IN SEMI-AUTOMATED VEHICLES- EXPERIMENTAL RESULTS

Chapter 8 suggested a decrease in driving performance after the automated phase accompanied by an inability of the driver to accurately assess their own fitness to drive. By way of a follow-up, Chapter 10 presented the results of the modelling of the psychophysiological and driving data to create prediction models. The aim of the models was to suggest a system of driver state monitoring to assure safe-take over and attention during the automated phase. It was possible to partially predict driver performance with linear equations based on the psychophysiological functions, which partially confirmed the hypotheses. A model predicting driving performance after take-over used factors derived from ECG, some factors derived from EDA, and scores from the NASA-TLX questionnaire. Although it was a statistically significant linear model, it explained only 22% of the variance in driving performance after the take-over. The model predicting attention during the automated phase used a factor derived from EOG and explained 23% of the variance in the attention test during automation. Both models presented valid associations between performance and physiology; however, at this stage, they are not ready to be implemented in road-going vehicles and requires more research.

1.4.10 CHAPTER 11: DISCUSSION AND CONCLUSIONS

Chapter 11 summarized all the work presented in the previous chapters, discussed it, and listed its limitations. It offered several recommendations for the manufacturers as well as directions for future research.

1.5 CONTRIBUTION OF KNOWLEDGE

This doctoral thesis presented a novel work on the psychophysiology of the driver in the semi-automated vehicles, circadian effect on semi-automated driving and driver state monitoring in the semi-automated vehicles. It supported previous concerns related to the underload of the driver during automation and its detrimental effect on the performance. Additionally, this research has provided evidence that such a result might be more exaggerated at night. Moreover, it has shown that drivers are not effective in self-assessment of their own fitness to drive. These were novel results because of the variety of psychophysiological methods used in this study simultaneously. Besides, two predictive models were proposed as systems of psychophysiological driver state monitoring. However, the predictive power of the models was low, so they are not ready to be implemented in the vehicles at this present moment. The models can be used as a basis for further psychophysiological research into driver state monitoring and the more general topic of the monitoring of any human operator of technology.

This thesis also offered a tool for the researchers to optimize laboratory work and select the most effective ways of noise reduction when using multiple physiological recording simultaneously at the laboratory. The novelty of this work is related to the high amount of physiological measures used simultaneously, the circadian approach to semi-automated driving, and predictive modelling of physiology in the semi-automated set-up.

1.6 FUTURE DIRECTIONS

This work presented the analysis of rich psychophysiological data gathered in the semi-automated driving experiment. The analysis presented in this thesis is only part of what could be undertaken with these data. When creating predictive models, only linear and binomial regression models were applied. Quadratic and exponential regression could be used to depict regularities in this dataset. Also, machine learning and deep learning algorithms could be used to analyse the physiology of the driver during automated mode and characterise patterns related to the performance decrease.

This work indicated drowsiness and sleepiness as detrimental for monitoring during automation and suboptimal mental workload and arousal as the most negative states for the take-over performance. These states should be further studied to better understand the driving risks. Especially, mathematical modelling of the association between mental workload, arousal, and performance could bring better insight into driver states and safety.

This experiment suggested that the monitoring role that driver must assume during semi-automated driving had a negative effect on driving safety. Therefore, it is recommended to keep drivers more involved in the active process of vehicle control until full automation is ready.

The results presented some circadian effects in the driver performance and state; however, the direction of the changes was not clear, and as so more research would be recommended in this area, especially in the forced-desynchrony protocol (Blatter & Cajochen, 2007).

2. RISKY DRIVER STATES IN SEMI-AUTOMATED VEHICLES- REVIEW

2.1 INTRODUCTION

In 2016 car accidents were the primary death reason for the people aged 15-29, and over 1.35 million people died in car accidents worldwide (World Health Organization, 2018). Based on the National Highway Traffic and Safety Administration (NHTSA) data, 94% of car accidents have been classified as being caused by human-machine system error (Melnicuk et al., 2016).

Automation offers the potential to mitigate or reduce such risk. It can also increase effectivity, alleviate workload and improve the transport capacity (Kyriakidis et al., 2019). In the aviation domain, the use of automation has risen rapidly over the past three decades and significantly improved safety (Chialastri, 2012). It is, therefore, anticipated that higher levels of automation could be incorporated into automobiles to reduce safety risks to road users and pedestrians alike (Kyriakidis et al., 2019).

However, there are different levels of automation that enable such requirements to different extents. The Society of Autonomous Engineers, as shown in figure 2.1, proposed a classification system ranging from 0 (no automation) to 5 (full automation). In the short term, full automation is not realistic due to legal challenges and technical limitations. In the coming decades, it is anticipated that levels 3 and 4 automation will become more prevalent. Therefore, these are the levels of automation that will form the basis of the current review.

As shown in figure 2.1, level 3 of the automation allows complete automated control over dynamic driving in particular circumstances, for example on the highway, but requires a human driver to stay fully attentive to the road and regain control over the dynamic driving functions (take-over) when requested. Level 4 of the automation allows fully automated control over dynamic driving in certain circumstances. This level of automation does not require a human driver to stay fully attentive to the road environment, but they still have to take-over when requested (Kyriakidis et al., 2019).

Table 2.1: A gradient of car automation levels based upon classification generated by Society of Autonomous Engineers (Kyriakidis et al., 2019).

Automation Level	Role of a human driver	Role of automation
0: Driver only	All aspects of the driving.	None.
1: Assisted automation	Driver carries on most of the driving tasks, except for one performed by automation.	The automated system performs either steering or acceleration/deceleration.
2: Partial automation	Driver carries on most of the driving tasks, except for the few performed by the automation.	One or more automated systems perform both steering and acceleration/deceleration.
3: Conditional automation	The driver carries driving in some periods of the time, and is expected to monitor the road, and respond when requested or in an emergency during the automated driving.	The automated systems perform all the driving tasks in the conditions for which they are designed if the human responds accurately to the requests to intervene.
4: High automation	The driver carries driving in some periods of time and should monitor the road but is not required to do so during automated driving.	The automated systems perform all the driving tasks in the conditions for which they are designed, even if a human does not respond accurately to requests to intervene.
5: Full automation	None.	The automated systems carry on all aspects of driving in all environments.

On 7th May 2016, a tragic accident of the Tesla Model S led to the death of the driver, Joshua Brown. Tesla Model S could be classified as an enhanced level 2 of automation, as it could automate both longitudinal and lateral aspects of driving. The computer vision system did not detect a white tractor on the background of the bright blue sky and drove straight under it. Unfortunately, the driver did not intervene and there were no recorded attempts by him to stop or redirect the vehicle. Also, for the majority of the drive, he did not have his hands on the wheel, despite the warning signals of the car (Banks et al., 2018). Some journalists claimed that he was watching a movie during the journey (Lambert, 2016; Neumann, 2016). The driver who was killed was actually an advocate for automation and was quoted saying that full attention is still required when using Tesla's autopilot (Lambert, 2016; Neumann, 2016). This crash has shown that the critical challenge is understanding how automation can best be integrated into current transport systems to maximise safety, comfort and productivity. The victim was a strong advocate of automation and knew that he was required to be attentive, but still, he was not. It exemplifies how an individual, even if they know about how behaviour should be governed, does not mean that they will or even have the capacity to act appropriately. Behaviour and knowledge are not always coherent.

Vigilance is an ability to maintain an attentive and alert state and is crucial for take-over safety. The problem is that people find it tiring, demanding and stressful to stay vigilant for more extended periods, especially during monotonous tasks. Whilst semi-automated driving requires vigilance; it also creates a challenging environment for its maintenance (Warm et al., 2008).

The transition between automated and human driving modes creates particular challenges, as does staying vigilant during prolonged automated driving. The period of automated driving can influence the driver's state, and at the moment of the transition, a driver might not be ready to take-over. Until level 5 automation is released in automobiles, there will always be a requirement for the driver to monitor the automation and at some point, to take back manual control. A critical question is whether the state of a driver can be monitored to assess if (1) they are monitoring the automation and (2) if they are in the right condition to safely transition from automated to manual control, commonly referred to as ‘taking back control' or ‘take-over’.

Understanding the physiology of the driver has the potential to provide insight into the mechanisms that underpin performance that could be defined as the ‘state' of the driver. In this work, a state was treated as a temporary psychophysiological condition of the driver (Chaplin et al., 1988). A critical task was defining the driver states that impact upon the capacity of the driver to optimally perform driving tasks, automation monitoring and, in particular, the transition between the two.

2.2 METHODS OF THE LITERATURE REVIEW

The databases used were Scopus, Web of Science, Google Scholar and DelphiS. Only the papers with the whole text available were used for the analysis.

To identify driver's states that present a threat for a take-over safety and to understand the capacity of the driver to regain control over the vehicle, the literature search was also extended to manual driving. It was motivated by two factors. Firstly, semi-automated driving is a relatively new area of research with a small number of experimental papers. At the same time, the risks that occur in the case of manual driving could also apply to this environment. Secondly, the task of the driver after the take-over is actual manual driving, so the risks that appear there are still valid for semi-automated driving, just modulated by new factors like prolonged time of inactivity during the automated phase.

The initial search terms used in the databases were: car OR driv* OR vehicle; human factors OR psycholog*; accident* OR disturb* OR fail*; state monitoring. Those terms were purposely selected to be broad and create a good base for the further iterative research in references and terms suggested by the initially identified literature, as recommended by the grounded theory (Heath & Cowley, 2004a). After primary search key references of the relevant publications were analysed to identify additional positions. The key references analysis was used iteratively. The initial number of results in the databases was: 53 in Web of Science, 178 000 in Google Scholar, 546 in DelphiS, and 211 in Scopus.

Exclusion Criteria: the articles that described risky traits like age, long-term health problems or personality were excluded, the same positions that analysed external risk factors, like bad weather. Articles that explored the risky states of the other types of operators, like pilots or controllers, were also excluded.

Inclusion Criteria: only full access articles in the English language were included in the criteria. Only the first 500 results of the search in the database results were included. The included materials analysed temporary driver's condition or a state as a risk factor for driving safety.

As a result, 168 papers were identified and included in the analysis of drivers' state. These papers were taken forward into the review in the next section.

2.3 RESULTS OF THE LITERATURE REVIEW

The aim of the review was to identify driver states that might present a threat to driving safety. Papers were evaluated to extract such factors and understand their role in driving safety. A list of all the identified papers, grouped by categories and provided definitions can be seen in Appendix 1. As shown in figure 2.1, drowsiness, fatigue, behavioural distraction, and cognitive distraction were the most frequently mentioned states that jeopardise the safety of the driving. The following chapters discussed them in details.

Some states beyond the cut-off point were also discussed because although they were not mentioned that frequently in the literature, there were other reasons to treat them as relevant for semi-automated driving. Sleepiness and sleep were included in the detailed descriptions because their definitions were very close to drowsiness, and they were often used as interchangeable terms (Johns, 2000). Suboptimal mental workload and insufficient situation awareness were frequently predicted as new types of risks related to semi-automated driving (Salmon et al., 2006; Young & Stanton, 2005), which made them relevant to the topic. Sleep inertia was mentioned only once; however, there were arguments provided advocating that it can become a new risk factor characteristic for semi-automated driving.

Literature provided a variety of definitions for each state. All the definitions were grouped into categories and presented in the table in Appendix 1. The subchapters provided only the most frequent definitions.

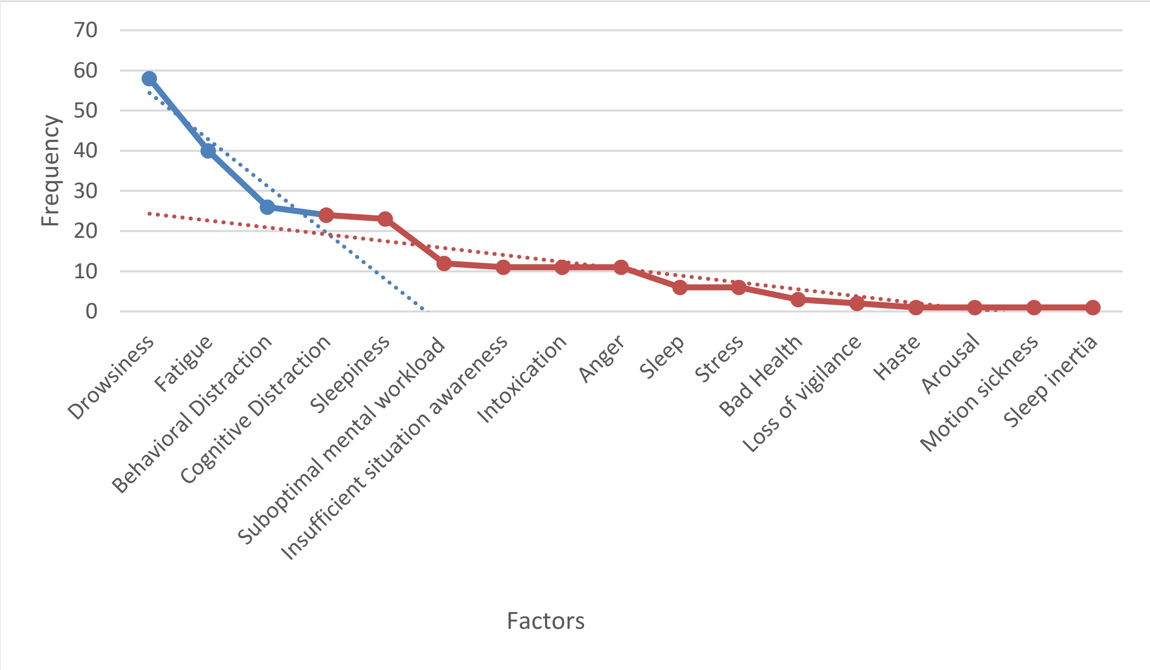

Figure 2.1: a scree plot presenting frequency of different factors in the analysed 168 papers. The blue and orange lines are the approximate trend lines for blue and orange parts of the plot. Their crossing point was a cut-off point. The trend lines were estimated rather than formally calculated.

2.3.1 DEFINITIONS

Literature provided an abundance of definitions for each of the analysed factors. Some of the definitions differed significantly from each other while other presented a very similar understanding of the concept and diverged only in small details. The definitions grouped into categories were presented in Appendix 1. Only one of the definitions was chosen for the detailed description based both on frequency and accuracy.

Exceptional cases were definitions of drowsiness, sleepiness and fatigue. The meaning of those terms varied depending on a scientific position. They were often used in the context of being tired and close to the state of sleep, but authors rarely defined them. Due to such ambiguity in the terminology, some scientists tried to come up with clear definitions.

Drowsiness was defined by Johns (2000) as a state between wakefulness and sleep determined with the use of the EEG oscillations pattern, eye movements, and muscle activity. It started before the onset of stage-1 sleep. Stage-1 sleep is an initial stage of sleep.

Sleepiness could be identified as a sleep propensity that is a probability of transitioning from wakefulness to drowsiness, and later into stage-1 sleep. Subjective sleepiness might be different from physical sleepiness and is related to mental and physical feelings associated with drowsiness (Johns, 2000).

Fatigue was defined as a subjective feeling of being tired due to the performed task. It was associated with a persistent need to stop it and it disappeared when a person switched to a different task (Johns, 2000). Fatigue could be divided into different categories: local physical fatigue, general physical fatigue, central nervous fatigue (which might be similar or equal to sleepiness), and mental fatigue (Dong et al., 2011).

While the term fatigue was mostly used according to the definition of Johns (2000) and Dong et al. (2011) sleepiness and drowsiness were often used interchangeably even within one paper. The same applied to the adjectives ‘drowsy’ and ‘sleepy’. They were rarely defined. One of the reasons might be that their outcomes on driving were very similar if not the same, and their physiological markers were also hard or sometimes impossible to distinguish (Johns, 2000). For these reasons, in this work, terms drowsiness and sleepiness were treated as synonyms and discussed collectively while the term fatigue was used in the understanding of Johns (2000) and Dong et al. (2011).

2.3.2 SLEEP

The factor of sleep was only identified in six papers; however, drowsiness that is a state in between wake and sleep was the most prevalent factor. As sleep and sleep drive play an essential role in drowsiness mechanisms (Johns, 2000), it is relevant to elaborate on it. Also, some papers did not use precise definitions and it was not clear if the accident risk they described was related to sleepiness or falling asleep.

Sleep is notoriously tricky to define. The Merriam-Webster dictionary gave a following definition of sleep: ‘the natural, easily reversible periodic state of many living things that is marked by the absence of wakefulness and by the loss of consciousness of one's surroundings, is accompanied by a typical body posture (such as lying down with the eyes closed), the occurrence of dreaming, and changes in brain activity and physiological functioning, is made up of cycles of non-REM sleep and REM sleep, and is usually considered essential to the restoration and recovery of vital bodily and mental functions' (Merriam-Webster, 2018).

Sleep is closely related to the sleep drive. The classical models of the sleep drive divide it into two components: a homeostatic and a circadian, with a correction for the sleep inertia (Akerstedt & Folkard, 1995). Sleep inertia is a state of the decreased cognitive and psychomotor functions and increased sleepiness just after waking up (Blatter & Cajochen, 2007). The term homeostatic refers to the balance between the amount of sleep, and the time that passed since waking-up. People who slept less, and have been awake longer have a stronger homeostatic drive to fall asleep, while people who slept a lot a short time ago have a weaker homeostatic sleep drive (Cluydts et al., 2002). The circadian drive refers to the body rhythms over the 24-hour cycle. People tend to feel more sleepy during the night and less sleepy during the day (Blatter & Cajochen, 2007). The newer models included the influence of the wake drive, also called arousal. The wake drive is related to the activity performed by the person, the body posture, and the potential soporific influences like boredom. The wake drive opposes the sleep drive, such that a highly aroused person might have troubles falling asleep in the night, even after a long period of wakefulness (Cluydts et al., 2002).

Sleep comprises of five stages characterised by different muscular and electroencephalographic activity (Ackermann & Rasch, 2014). In the case of driving, people mostly experience NREM stage-1, which is the entry-stage of sleep (Ackermann & Rasch, 2014).

Horne and Reyner (1995) reported that sleep-related accidents comprised 16% of the overall accidents and even 20% of motorway accidents. They suggested that this number might be understated. They observed that the sleep-related accidents occurred most often between midnight and 2:59 a.m. and had two peaks during the day, around 6 a.m. and 4 p.m. They emphasised that sleep-related accidents led to more serious injuries. Hakkaken and Summala (2001) who studied truck accidents, found that 5.3% of truck drivers who caused the accidents fell asleep behind the wheel, even though none of them had increased daytime sleepiness. Sahayandhas et al. (2013) cited the National Sleep Foundation, pointing out that 28% of Americans fell asleep behind the wheel. Higgins et al. (2017) in their review gave even larger statistics stating that up to 41% of American drivers reported that they fell asleep behind the wheel at least once.

In their review paper, Horne and Reyner (1999) stated that the circadian factors are as crucial for falling asleep behind the wheel as time spent driving before the incident. They pointed out that many drivers were not aware of their sleep episodes during driving, which could lead to a worse sleep control, but also a statistical underestimation of the number of accidents caused by a driver falling asleep. Moreover, even when drivers were aware of increasing sleepiness, methods used by them to alleviate it were mostly ineffective.

The group that was at the most significant risk of falling asleep behind the wheel were young, educated men, who drive a lot, with obstructive sleep apnea syndrome. Falling asleep was also most likely when they are driving alone (Gonçalves et al., 2015; Sagberg, 1999).

Sleep was closely related to drowsiness, sleepiness, and sleep inertia. Individuals who experienced sleep deficits felt more drowsy. At the same time, people experiencing drowsiness fell asleep much quicker (Higgins et al., 2017), while sufficient sleep decreased the probability of dozing-off (Cluydts et al., 2002), except for the period of sleep inertia (Blatter & Cajochen, 2007).

2.3.3 DROWSINESS AND SLEEPINESS

Drowsiness was identified as the driving safety risk factors in fifty-eight papers, while sleepiness in twenty-three. As explained in the Definitions subchapter, for the purpose of this work, those two terms were treated as synonyms. Drowsiness and sleepiness were then the most prevalent factors as in sum they were mentioned in eighty-one identified articles.

Drowsiness and sleepiness are the terms describing a state in between sleep and wake that is associated with reduced performance and attention (Jackson et al., 2016).

Various authors indicated them as a high and underestimated risk for driving safety and stressed the need of developing proper countermeasures (Jackson, Raj, et al., 2016; Smith et al., 2003; Van Winsum, 2000). Statistical estimations presenting the prevalence of accident caused by drowsiness varied depending on the methodology and place of the research. Depending on a source, up to 75% of drivers reported driving while feeling drowsy, up to 45% of accidents and up to 40% of fatal accidents were related to driver's drowsiness (Bekiaris, 1999; Ebrahim et al., 2013; Kwai et al., 2016; Ma'touq et al., 2014; Murata et al., 2017.; Rodríguez-Ibáñez et al., 2011; Solaz et al., 2016; Wang et al., 2017). Driving while drowsy increased the risk of an accident fivefold (Wang et al., 2016). Fairclough and Graham (1999) found that severe drowsiness created a hazard for driving safety as high as alcohol intoxication.

According to Higgins et al. (2017), drowsiness behind the wheel could lead to the accident in two ways: driver might either shift from drowsiness to sleep and drive off from their lane or stay in the state of drowsiness and cause an accident due to their decreased cognitive capabilities. Drowsy drivers had impaired visual perception, sustained attention, lower cognitive functions, decision making, reaction time (da Silveira et al., 2016; Higgins et al., 2017; Jackson, Kennedy, et al., 2016; Kumari & Kumar, 2017), volition (Johns, 2000; Yang & Jeong, 2015) visual sensitivity, late motor processing, speed and accuracy of various cognitive processes, working memory, short-term memory, executive functions, supervisory control, spatial orientation, situation awareness, mathematical processing, motor task abilities and divergent thinking capacity (Ftouni et al., 2013; Howard et al., 2014; Krajewski, Batliner, et al., 2009). Additionally, drowsiness was often observed to decrease mood (Krajewski, Batliner, et al., 2009). Drowsiness increased the risk of lane departure (Liu et al., 2009; Sahayadhas et al., 2013) and speed variability (Sahayadhas et al., 2013); however, effects of drowsiness were not always visible in the driving performance, and some drivers could perform well despite high drowsiness (Sahayadhas et al., 2013).

Drowsiness-related accidents had their peak at night, early morning hours, and a smaller rise in the afternoon hours. It was more prevalent at night, but people could experience high sleepiness also during the day. Goncalves et al. (2015) reported that in Europe, the prevalence of high daytime sleepiness among drivers varied from 1% to 8%. Excessive daytime sleepiness was associated with an increased frequency of falling asleep behind the wheel and an increased amount of sleepiness-related accidents.

There were mixed findings related to the drivers' awareness of their own sleepiness. Some authors reported that drivers did not know that they were going to fall asleep (Filtness et al., 2017; Ftouni et al., 2013; Howard et al., 2014), while other that drivers manifested relatively high awareness of their sleepiness level (Akerstedt et al., 2013; Filtness et al., 2017).

Drowsiness could be addressed mostly to the four factors: circadian phase, amount of time spent without sleep (Johns, 2000; Kumari & Kumar, 2017; Rahman et al., 2015; Sahayadhas et al., 2015), time on task (Sahayadhas et al., 2013; Wang et al., 2017), and a level of arousal (Rahman et al., 2015). The drivers who were reported to be most vulnerable to sleepiness were the young men who were alone in the car. Moreover, professional drivers, shift workers, and people working extended hours were more likely to experience drowsiness while driving (Higgins et al., 2017; Johns, 2000; Wierwille et al., 1994).

2.3.4 FATIGUE

The risk factor of fatigue was identified in forty papers.

Fatigue is a state caused by the prolonged performance of one task. It can comprise of feeling tired, sleepy, need to stop the task, and decreased cognitive or muscular performance (Johns, 2000). It can be easily distinguished from drowsiness, because a period of rest or change of the task alleviates fatigue, while it often makes drowsiness worse. Fatigue is also steady, cumulative process increasing with the time on task, while drowsiness can fluctuate rapidly (Borghini et al., 2014).

Two types of fatigue could be distinguished based on the causal factor, sleep-related fatigue and task-related fatigue. Sleep-related fatigue could be caused by both insufficient sleep and driving during the circadian night. Task-related fatigue could be caused by a task that requires too much workload as well as by a monotonous task that requires only sustained attention (May & Baldwin, 2009).

Many car accidents could be addressed to fatigue (Borghini et al., 2014; Di Stasi et al., 2015; Haq & Hasan, 2016; Lal & Craig, 2002). It decreased driving performance (Haq & Hasan, 2016; Lal & Craig, 2001; Melnicuk et al., 2016) and drivers themselves indicated it as a severe driving risk (Häkkänen & Summala, 2001). Depending on the publication up to 45% of the overall car accidents and 30% of the fatal road accidents were addressed to fatigue (Fan et al., 2010; Fu et al., 2016; Simon et al., 2011). Even up to 55% of the drivers have driven while fatigued over a period of a year (Wijesuriya et al., 2007).

Symptoms of the fatigue in drivers were body pains and discomfort, drowsiness, decreased mood, slower activity, irritability, attention deficits, problems with signs observation, decreased performance, and difficulties in decision making (Haq & Hasan, 2016; Li et al., 2015; Tran et al., 2014). Fatigue increased reaction time, but what was interesting it decreased reaction time in case of wrong responses (Milosevic, 1997). Fatigue also caused a decrease in attention (Dhupati et al., 2010; Lal & Craig, 2001; Liu et al., 2010), could lead to falling asleep, a decrease in road position control and speed control (Lal & Craig, 2002, 2005; Mittal et al., 2016), increased subjective stress and workload, increased heading error, and reduced steering activity (Matthews & Desmond, 2002). Even a moderate level of fatigue could induce driving mistakes comparable to ones caused by alcohol intoxication (Arnedt et al., 2001; Jap et al., 2009).

Fatigue was associated with sleepiness, and the peak of the fatigue-related accidents was synchronised with the circadian phases with the highest sleepiness level (night and afternoon) (Arnedt et al., 2001; Lal & Craig, 2005; Puspasari et al., 2015). The long, monotonous or boring tasks also increased fatigue (Borghini et al., 2014; Di Stasi et al., 2015; Lal & Craig, 2001), as well as the highly demanding tasks (Heikoop et al., 2016). There was a U shaped association between fatigue and a mental workload, too low and too high a mental workload could induce fatigue (Heikoop et al., 2016).

Drivers more susceptible to fatigue were younger (Häkkänen & Summala, 2001), and had a disturbed sleep cycle (Lal & Craig, 2001). The group significantly affected by fatigue were professional drivers and they frequently addressed their accidents to this factor (Borghini et al., 2014; Lal & Craig, 2001).

A subjective feeling of fatigue was observed to often dissociate from the fatigue-related performance drop (Wijesuriya et al., 2007). Drivers reporting feeling fatigued not always presented performance drop, and fatigue-related performance drop was not always accompanied by a subjective feeling of fatigue (Brown, 1994). Similar dissociation between subjective state and performance was also observed in the case of drowsiness (Filtness et al., 2014; Ftouni et al., 2013). Such a lack of awareness might increase the risk of fatigue-related accidents.

2.3.5 BEHAVIOURAL DISTRACTION

Twenty-six identified papers described behavioural distraction as a driving safety risk factor.

In this work, driver's distraction was divided into two categories, behavioural- associated with some physical activities unrelated to driving, like scrolling or texting, and cognitive-associated with the mind-wandering, attending to something else than the road or not paying enough attention to the driving tasks.

The term ‘behavioural distraction' was rarely defined, often replaced by a description of the distracting behaviours. The examples of mentioned distracting behaviours were, talking on the mobile phone (Caird, 2015; Márquez et al., 2015; Seiler, 2015), dialling (Klauer et al., 2014; Petridou & Moustaki, 2000) or lighting a cigarette (Petridou & Moustaki, 2000), but the list is as long as a human's imagination (Caird, 2015; Klauer et al., 2014; Petridou & Moustaki, 2000). Behavioural distraction can be defined as using attention for the activity competing with the primary task (Hosking et al., 2009) and was sometimes called ‘eyes-off-road’ (Liang & Lee, 2010).

Disturbed attention was often mentioned as an essential risk factor for driving (Caird, 2015; Chan et al., 2016; Rumschlag et al., 2015). According to the multiple resources theory when the different thoughts and actions compete for the attentional resources the performance decreases (Liang & Lee, 2010), hence various activities undertook while driving might decrease a driving performance.

Talking on the phone while driving increased the risk of the car accident even four times (Redelmeier & Tibshirani, 1997), decreased the lateral position control, delayed the speed adaptation (Lamble et al., 1999), made the steering behaviour more violent within a city (Brookhuis et al., 1991), increased the number of the off-road excursions (Haigney et al., 2000), increased the risk of the traffic lights missing (Strayer & Drew, 2004), increased the reaction time (Horrey & Wickens, 2006), and increased the mental workload (Brookhuis et al., 1991). Some authors claimed, that a decrease in the performance was worse with the hand-held phones than with the hands-free phones (Haigney et al., 2000), while some argued that the outcome of those two is the same (Horrey & Wickens, 2006; Márquez et al., 2015; Strayer & Drew, 2004). Individuals with a tendency to the compulsive phone using, and those that were awaiting a phone call were more likely to talk on the phone while driving (O'Connor et al., 2017).

Texting while driving was also a prevalent behaviour and created a significant safety risk estimated as even more prominent than talking on the mobile phone (Drews et al., 2009; Hosking et al., 2009). Texting while driving impaired lateral vehicle position control (Rumschlag et al., 2015), traffic signs recognition, road focus, reaction time (Hosking et al., 2009), and lane maintenance (Drews et al., 2009; Hosking et al., 2009). However, texting drivers also decreased their speed (Hosking et al., 2009), and kept a bigger distance to the vehicle in front, as a compensation attempt (Drews et al., 2009). People who had internet access on their phones, talked on the phone while driving, were in the car with other drivers who text and drive, as also people who sexted regularly and used their mobiles when they were bored were more likely to text behind the wheel (Seiler, 2015).

Hoel et al. (2011) studied the effect of the different types of behavioural distraction on driving safety. They found that it mostly led to detection and execution failures, but unlike other studies, they concluded that it rarely led to accidents. Klauer et al. (2014) found that most of the behavioural distractors created more risk for the novice than for the experienced drivers. Talking on the phone did not increase the risk of the crash, while dialling on the phone did. The effects of a distraction on driving performance were related to working memory and working memory capacity. Individuals with a bigger working memory capacity performed better under distracting stimulation (Ross et al., 2014). Also, experienced individuals (Pope et al., 2017) and individuals in middle age were less susceptible to the distraction-related performance drop than young, older, and novice drivers (Rumschlag et al., 2015). At the same time young, and middle-aged drivers were more likely to engage in distracting behaviours than older ones (Pope et al., 2017). Behavioural distraction was a very prevalent phenomenon, and most of the drivers involved in the distracting activities from time to time (Márquez et al., 2015; Pope et al., 2017; Seiler, 2015).

2.3.6 COGNITIVE DISTRACTION

Twenty-four of the analysed papers identified cognitive distraction as a driving safety risk factor.

Cognitive distraction could be defined as the diversion of attention towards mental activity unrelated to the main task (Wesley et al., 2010), also called ‘mind-off-road’ (Liang & Lee, 2010). Different expressions were used in the literature in the same meaning as a cognitive distraction: attention laps (Parker et al., 1995), diminished vigilance (Ji & Yang, 2002), inattention (Bando & Nozawa, 2015; Casner et al., 2016; Regan et al., 2011), mind wandering (He et al., 2011), lowered concentration (Kawanaka et al., 2013) and hypovigilance (as a term including both cognitive distraction and drowsiness) (Sahayadhas et al., 2015).

Driver’s cognitive distraction was often described as a cause of car accidents (Melnicuk et al., 2016; Sahayadhas et al., 2015; Yang & Jeong, 2015); however, depending on the study there was a different estimation of the prevalence. Research by Parker et al. (1995) has shown that only a small amount of car accidents was caused by lapses in attention. There were also other studies proving that behavioural distraction is much more detrimental for driving than cognitive (Hoel et al., 2011). However, many other authors listed it as a leading cause of the accidents estimating that even up to 80% of crashes can be related to inattention (Melnicuk et al., 2016; Miyaji et al., 2009; Parnell et al., 2016). Some studies classified even 40% of the driver's errors as related to inattention (Stanton & Salmon, 2009).

Inattention can have even larger negative effect on driving of semi-automated vehicles, due to the periods of automated driving, and was identified as a risk factor specific for semi-autonomous driving (Casner et al., 2016; Dogan et al., 2017; Heikoop et al., 2016; Merat et al., 2014).

Cognitive distraction affected the quality of driving by reduction of attentive resources available for driving tasks. According to the multiple resources theory, when different thoughts compete for the attentional resources, it decreases the performance (Liang & Lee, 2010). Cognitively distracted drivers often concentrated their gaze on the centre of the road but had reduced peripheral vision and limited ability to detect targets from the whole range of the visual field (He et al., 2011; Liang & Lee, 2010). They also had a longer reaction time (He et al., 2011; Yang & Jeong, 2015), worse signal detection, vigilance, and memory (He et al., 2011), longer time on task, and more steering errors (Yang & Jeong, 2015). They manifested frequent detection errors, diagnosis failures and prognosis failures (Hoel et al., 2011). Cognitive distraction also led to recognition errors (Melnicuk et al., 2016).

Cognitive distraction could be caused by a variety of factors, like overfamiliarity with the environment, long driving experience, generally low level of attention, thoughts and concerns (Hoel et al., 2011), fatigue or drowsiness (Chakraborty & Aoyon, 2014). According to PARRC model by Parnell et al. (2016) main factors contributing to driver's distraction were goal conflict, adaptation to the demands, behavioural regulation, goal priority, and constraints of the resources. There is an inverted- U relationship between attention and mental workload, too small and too big mental workload was associated with attention deficits. Stress was shown to decrease attention (Heikoop et al., 2016). A propensity for cognitive distraction behind the wheel increased with age (Stanton & Salmon, 2009).

In the driving errors, taxonomy by Donald Norman, errors related to the lack of attention were categorised as slips that led to unintended wrong actions. The examples of slips were misperception, action intrusion or omission of action. Other studies showed that cognitive distraction could lead to recognition errors (Stanton & Salmon, 2009).

2.3.7 SUBOPTIMAL MENTAL WORKLOAD

Only twelve of the analysed papers listed suboptimal mental workload as a risk factor for driving safety. However, it was described as a distinctive risk factor that will probably increase with the development of car automation, due to the small number of tasks in the automated mode. Drivers inactive during the longer periods of automated driving might experience underload and related to that cognitive decrease (De Winter et al., 2014; Heikoop et al., 2016; Melnicuk et al., 2016). Because it is a problem that might become more significant with the development of automation, this factor was also included in the detailed description of the risky driver's states.

The mental workload was related to the proportion of available mental resources of the operator and the number of resources necessary for the task. If the resources required by the task exceeded an optimal level, the workload was too high, while if they were lower than the optimal level, the workload was too low. However, the level of optimal mental workload was individual and depended on circumstances, task and operators skills (Palinko et al., 2010; Young & Stanton, 2005).

In driving, the mental workload could be affected by the driver's skills, age and state, road and traffic-related circumstances as well as a car (Young & Stanton, 2005).

The mental workload was also sometimes addressed using different terms such as mental effort (Brookhuis & de Waard, 2010), cognitive load (Engström et al., 2017; Palinko et al., 2010), visual perception, mental/cognitive processing, overload/underload, cognitive resources or cognitive activity/task (Heikoop et al., 2016).

Gregerseb and Bjurulf (1996) analysed the reasons for accidents in young drivers, concluding that their lack of experience increased mental workload making it more difficult to drive safely. There were also other studies supporting the negative effect of an overload on performance (Brookhuis & de Waard, 2010; Melnicuk et al., 2016; Yang & Jeong, 2015).

The too-high mental workload was detrimental to the driving performance, but it came out that too low a workload was at least as bad. The association between mental workload and performance can be represented with an inverted U-shaped function, with low performance associated with low and high mental workload, and the most optimal performance with the middle workload (Heikoop et al., 2016; Young & Stanton, 2002). The association between the mental workload and attention was similar. In contrast, the association between mental workload and stress could be represented by a U-shaped function, with the low and high levels of the mental workload associated with larger stress than the middle levels (Heikoop et al., 2016).

One of the explanations of the negative underload effect on the performance was that the operator tends to adjust the level of used resources to the situation difficulties, hence uses fewer resources in the case with lower demands (Young & Stanton, 2002). Engstrom et al. (2017) suggested that the increased mental workload affected only those aspects of the driving performance that relied on the cognitive control, while did not affect functions that were habituated (Engström et al., 2017). Various studies showed that mental workload was increased by stress, task demands, and attention. Feedback during tasks had the potential both to increase mental workload because of information amount and decrease the mental workload because of increased situation awareness (Engström et al., 2017; Heikoop et al., 2016).

Automation was presented to decrease the mental workload (De Winter et al., 2014; Stanton et al., 1997); however, Yong and Stanton (2002) suggested that it can also sometimes increase mental workload due to the complication of the computer system, and the abundance of the system modes. An example can be a category of operating errors addressed as a mode error when the operator finds it hard to realise what mode is the system in, and because of this undertakes actions unsuitable for the situation (Stanton & Salmon, 2009). It needs to be borne in mind, that not every level of mental workload increase, or decrease can be observed in the performance. The changes in performance also depended on the level of available cognitive resources. If such a level is high, an increase of mental workload might not change the performance but decrease a potential reactivity to the additional tasks (Parasuraman et al., 2008; Young & Stanton, 2005).

2.3.8 INSUFFICIENT SITUATION AWARENESS

The risk factor of insufficient situation awareness was only mentioned in the eleven analysed positions. Still, the same as in the case of mental workload, some authors stressed that it is a characteristic risk for automated driving. The reasons for this are feedback given by the system, the passive role of the operator, and the necessity of the long-term sustained attention, that humans are poor at (Endsley, 1996; Heikoop et al., 2016; Stanton & Young, 1998).

Endsley (1996) defined situation awareness as ‘the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future. She divided situation awareness into three levels: level 1 based on the perception of the situation, level 2 based on the correct interpretation of the perceived stimuli, and level 3 based on the effective prediction of the near-future. The deficit in any of those three levels defined an insufficient situation awareness. Other authors simply describe situation awareness as ‘knowing what is going on' and argued if Endsley’s definition includes the full scope of the topic (Stanton & Salmon, 2009).

A low level of situation awareness might lead to bad decision making and result in the accident (Borghini et al., 2014; Endsley, 1996). According to Reason's classification of the driving errors, mistakes mostly come from insufficient situation awareness. Many other driving errors fully or partially came from the lack of situation awareness, and even up to 88% of the car accident were addressed to this risk factor (Stanton & Salmon, 2009).

An interesting example of insufficient situation awareness in automation is a mode error, which happens when the operator does not realise what mode is the machine in (for example manual or automated driving modes) and chooses wrong actions based on incorrect assumptions. The abovementioned issue is related to the broader problem of feedback. If the driver does not receive proper feedback from the automated car might get confused and detached from process monitoring. Understanding feedback also requires a general knowledge of the automation that might present a challenge to some drivers (Endsley, 1996). However, automation might confuse but could also lead to increased situation awareness due to decreased mental workload. But yet, workload only reduced if the driver stayed focused on the main task. If they engaged in competing activities, situation awareness tended to decrease even more (De Winter et al., 2014).

Another listed challenge of automation was that it could shift the driver's tasks from manual control to monitoring. Monitoring requires sustained attention and vigilance that are difficult to maintain for many people. Especially drivers with a high level of trust towards automation might detach from their monitoring task (Dogan et al., 2017). The situation awareness might also be decreased by necessity of long-term machine monitoring, too big reliance on the automation, false alarm incidents, passive role of the operator (Stanton & Young, 1998), inappropriate feedback (Heikoop et al., 2016), lack of automation understanding, the complexity of the automated systems (Endsley, 1996), distraction (Dogan et al., 2017; Johannsdottir & Herdman, 2010), too big or too small mental workload (Borghini et al., 2014; Heikoop et al., 2016), stress, and too big trust to the automation (Heikoop et al., 2016). The operators were also shown to have smaller situation awareness when sleep deprived (Sneddon et al., 2013).

Situation awareness could be improved by an optimal mental workload and attention (Heikoop et al., 2016). Some studies showed that some levels of automated support could increase the driver's situation awareness, but only when a driver was motivated, properly instructed, and not engaged in the competing task. Unfortunately, drivers using vehicles with higher levels of automation were strongly inclined to involve in activities not related to driving (De Winter et al., 2014).

2.3.9 SLEEP INERTIA

Sleep inertia was only investigated in one identified paper. However, this work argued that it might become a new type of risk in semi-automated driving, also the identified article researched sleep inertia in the context of semi-automated driving and take-over after waking-up.

In the lower levels of automation falling asleep already creates such a high risk, that its aftermaths following the waking-up moment are not even considered. However, in the higher levels of the automation driver might sleep during the automated mode without causing an accident and then wake-up to take-over. In such a case, sleep could influence the driver's performance after waking-up, not at the moment of falling asleep.

Sleep inertia is a state of increased sleepiness, hypovigilance and decreased mental, and physical functions after waking up (Ferrara & De Gennaro, 2000). It could be considered a paradoxical state because people experience high sleep drive when homeostatic sleep drive should be absent. Some authors hypothesized that that mechanism of sleep inertia is a low level of arousal caused by prior sleep (Tassi et al., 2006), other stress that it is not easy to distinguish between cognitive and psychomotor effects of sleep inertia and sleepiness. It might be that they can be treated as the same state, just caused by the different factors (need for sleep vs effect of sleep) (Balkin & Badia, 1988).

Cognitive functions that were mainly affected by sleep inertia were the ones that require high accuracy, attention (Ferrara & De Gennaro, 2000), decision making (Bruck & Pisani, 1999) and speed (Hofer-Tinguely et al., 2005). Even though, these functions were affected the most, the effect of sleep inertia could also be observed in working memory, grip strength, steadiness and coordination, time perception, complex behaviours, logical reasoning, arithmetical operations and many other (Tassi & Muzet, 2000).

The length of sleep inertia could vary from one minute up to four hours, but the most intensive symptoms occurred between five to twenty minutes (Kolff et al., 2003). Not all the functions were restored at the same time, and some of the mental processes came back to normal after no more than four minutes (Tassi & Muzet, 2000).

Sleep inertia might be stronger and last longer in a case when a person woke-up from a sleep that did not last sufficiently long (Kolff et al., 2003; Tassi et al., 2006). The strength and length of sleep inertia were modulated by a circadian phase, a length of the sleep, and a level of sleep deprivation (Ferrara & De Gennaro, 2000; Muzet et al., 1995). Longer and stronger sleep inertia occurred mostly in the night and after sleep deprivation (Tassi & Muzet, 2000). Experimental research suggests that the early stage of sleep inertia might be more influenced by the depth of preceding sleep while following stages by the circadian phase (Wilkinson & Stretton, 1971). The stage of sleep prior to awaking also influenced the strength of sleep inertia (Tassi & Muzet, 2000).

Driving performance highly relies on attention (Giorgetti et al., 2015; McKenna, 1998) and speed of information processing, likewise on many other cognitive functions (Brouwer &

Withaar, 1997). Because of that, all the states distorting attention or speed might also distort driving safety. Ferrara and De Gennaro (2000) suggested that sleep inertia is an important reason to restrain from naps when an operator might be required to perform the demanding tasks just after waking up. Also, Muzet et al. (1995) indicated sleep inertia as an important side effect of napping at work and suggested that operators after the nap should be assigned to the tasks that are less affected by sleep inertia. While it might be a good strategy for the industry, there is no diversity of tasks in manual driving. It should also be taken into account that naps during the automated driving mode might be related to prior sleep deprivation or night driving. Hence effects of sleep inertia could be especially strong. The study on sleep inertia in semi-automated driving and take-over showed that reaction time increased and the quality of take-over significantly deteriorated when participants were in the state of sleep inertia (Wörle et al., 2020).

2.3 CONCLUSIONS

This chapter reviewed the literature to identify driver's states that could present a risk for semi-automated driving. It identified sleep, sleepiness, fatigue, behavioural distraction, cognitive distraction, suboptimal mental workload, insufficient situation awareness, and sleep inertia. Research showed that drivers were often unaware of these states (Ashleigh J Filtness et al., 2014; Wijesuriya et al., 2007), which creates a need for a driver state monitoring system. As so, chapter 3 reviewed available tools of psychophysiological monitoring of the driver state. The big challenge for such devices is that driver's state and performance are not perfectly correlated and a decreased driver state might not always lead to a reduced performance if the skills or mental capacity of the driver were high (Ross et al., 2014).

3. METHODS OF THE DRIVER STATE MONITORING- REVIEW

3.1 INTRODUCTION

In Chapter 2 a number of states that created a risk for driving safety were identified: drowsiness/sleepiness, fatigue, behavioural distraction, cognitive distraction, suboptimal mental workload, insufficient situation awareness, sleep, and sleep inertia. Studies showed that drivers are often not accurate in the assessment of their state (Filtness et al., 2017; Ftouni et al., 2013; Howard et al., 2014); therefore, methods that could detect risky states could increase driving safety. There have been many attempts to create systems of driver's state monitoring (Aghaei et al., 2016; Bekiaris, 1999; Parasuraman, 2011). This chapter aimed to review available measures of the driver's state.

Available reviews of the driver's state monitoring mostly concentrated on monitoring of one state (Lal & Craig, 2002) and did not present a full scope of available methods. The review is meant to play the role of a guide over known measures of the driver's state.

3.2 METHODOLOGY OF THE LITERATURE SEARCH

The search was initially conducted in Web of Science, Scopus, Google Scholar and DelphiS. The initial search terms were combined from the name of one of the risky driver's states identified in Chapter 2 for example sleep and following terms measur* OR indic* OR detecti*; psychophysiolog* OR physiolog* or psycholog*. After identification of several methods of driver state monitoring the second search was conducted. During the second search risky driver's states were combined with names of different psychophysiological methods to obtain detailed knowledge about particular measures, for example: sleep; EEG, measur* OR indic* OR detecti*. A further search was based on the relevant citations from the analysed papers, as recommended in the grounded theory (Heath & Cowley, 2004).

Inclusion criteria. To be included in the analysis, the paper had to be in English with full text available. Both experimental articles and reviews were included. The article had to describe a psychophysiological measure of the driver's state or psychophysiological measurement that can detect one of the states risky driver states (see chapter 2). Only papers describing in details indicators used to detect the state were included in the further analysis.

Exclusion criteria. Studies conducted on animals were excluded from the analysis. Also, studies that investigated diagnosis with the use of psychophysiological methods on the pathological health states were not analysed. Papers that described a technique of state detection but did not include indicators of the state were excluded from further analysis. For example, an article that stated that EEG can be used to detect drowsiness but did not state what frequencies can gauge it was excluded from the analysis.

As a result, 136 papers were included in further analysis.

3.3 RESULTS OF THE LITERATURE REVIEW

One hundred and thirty-six papers were included in the analysis of psychophysiological methods. They were either experimental studies or literature reviews exploring measures of the driver's state or measures of one of the states previously identified as driving safety risk (see Chapter 2) but in a different context than driving.

Out of one hundred-thirty-six identified papers, thirty-two explored EEG with an analysis of oscillations, twenty hybrids of measures, twenty eye-tracking, eighteen ECG, sixteen EOG, nine fNIRS, nine EDA, seven acoustic speech analysis, seven ERP, six EMG, six questionnaires, six blood pressure, five infrared video camera, four facial expression, four saliva analysis, four body temperature, three pupillometry, three respiration, two driving performance, two psychomotor performance, one head movements, one oximetry, one actigraphy, one blood glucose, and one Doppler flow meter. The detailed list of all the analysed factors, measured states and references was included in the table in Appendix 2. As shown in figure 3.1, an arbitrary cut-off point was placed after the second change of slope representing the frequency of different psychophysiological measures in the literature. A cut-off point allowed to select five measurement methods for further analysis: EEG, hybrid of methods, eye-tracking, ECG and EOG. Their meaning for driving research and safety was described in the separate chapters.

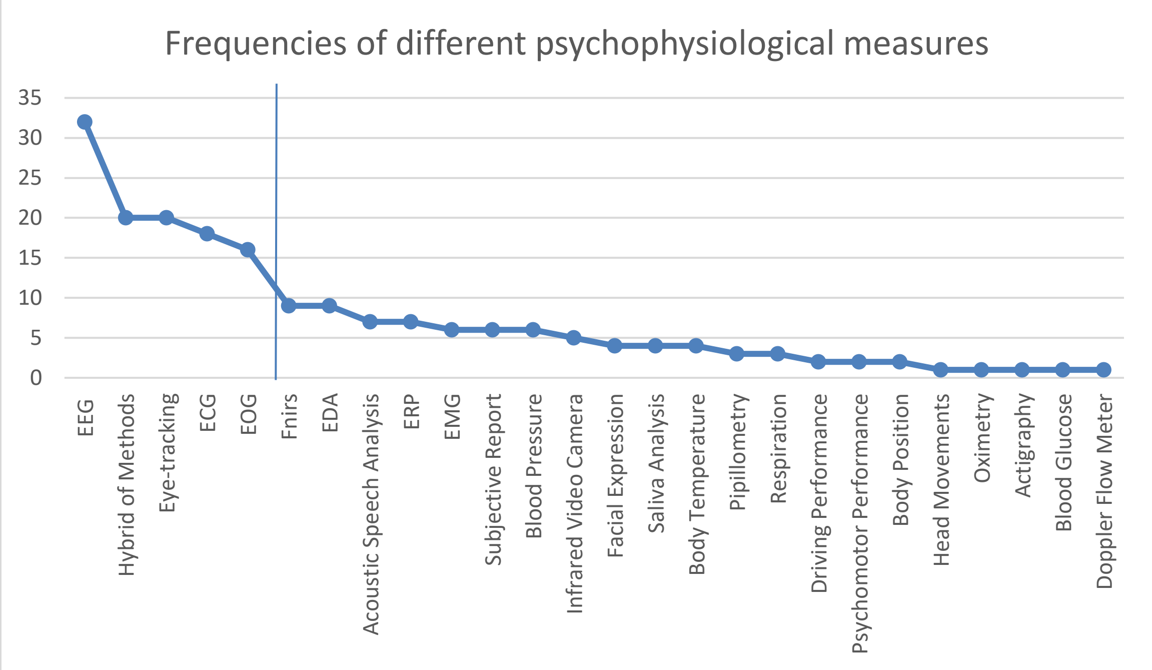

Figure 3.1: A scree plot representing a frequency of different psychophysiological measurements within the identified literature. A blue, vertical line represents a cut-off point. The cut-off point was chosen arbitrarily rather than based on the trend lines to stay more inclusive.

3.3.1 ELECTROENCEPHALOGRAPHY

EEG is a measure of the inhibitory and excitatory postsynaptic potentials of the cortical nerve cells. It can be either used to analyse individual potential (ERP) or frequencies of potential oscillations, widely just called EEG.

Frequencies of oscillations are often grouped into categories: Delta 0.5-4 Hz, theta 4-7 Hz, alpha 8-13 Hz, beta 13-30 Hz and gamma over 30 Hz. Different frequencies over different brain regions are related to different mental states and activities (Lal & Craig, 2001). Because of this, the measurement of the oscillations can be used for state detection.

In the identified literature, EEG was used to detect drowsiness, fatigue, vigilance decrease, mental workload, behavioural distraction and episodes of sleep. Thirty-two papers reported the use of EEG as a state detection method. The detailed list of algorithms used to detect those states was presented in the table in Appendix 2, and this chapter will only give a brief description.

In a majority of papers, the analysis was based either on power in one or more frequencies (Dhupati et al., 2010) or spectral entropy (Sriraam et al., 2016). However, some studies used machine learning algorithms like an artificial neural network (ANN) (Subasi, 2005) or a support vector machine (SVM) (Hogervorst, Brouwer, & van Erp, 2014) and found multiple features for classification. Some authors used devices with a state classification already implemented by the company, like B-Alert (Berka et al., 2004). Interesting exceptions were experiments that used EEG device to detect eye-movements, not brain activity (Jiao & Lu, 2016).

Drowsiness, fatigue, decreased vigilance and sleep were often related to power increase in slower frequencies, like theta and delta and a decrease in faster frequencies like beta (Borghini et al., 2014). Alpha frequencies were controversial because they were both reported to increase (Dhupati et al., 2010) and decrease (Oken et al., 2006). Increased alpha power could be treated as an indicator of mental activation as well as drowsiness, but only when accompanied by slow eye movements. Therefore, in the case of alpha measures, eye movements should be taken into account (Jödicke et al., 2013). The increased mental workload was repeatedly reported to be correlated with lower frequencies, suggesting cognitive activation decrease (Borghini et al., 2014; Kamzanova et al., 2014). Borghini et al. (2014) reported that behavioural distraction could be detected with an increase of theta power.

Summarizing, most of the methods developed so far detected risky driver's state with an increase in slower frequencies power, slow eye movements or machine learning algorithms.

3.3.2 HYBRIDS OF MEASURES

Twenty papers reported the classification of states based on a variety of psychophysiological measures. They used combinations of methods and classified them with various machine learning algorithms. Some of them reached very high classification accuracy, for example, 99.3% (Yeo et al., 2009). However, it has to be taken into account that machine learning equations were separately trained for every user and based their classification only on the pattern observed during the training phase (Alpaydin, 2014).

Papers using hybrids of methods described attempts to detect drowsiness (Ha & Yoo, 2016), cognitive distraction (Miyaji et al., 2009), stress (Healey & Picard, 2005), mental workload (Bundele & Banerjee, 2010) and fatigue (Wilson & Russell, 2003). All the algorithms used at least two psychophysiological measures and extracted multiple features from them. The details of the algorithms and psychophysiological measures were presented in the table in Appendix 2.

3.3.3 EYE-TRACKING

Twenty identified papers used eye-tracking as a state measure. Eye-tracking devices used video cameras to record ocular behaviours. The analysed factors were changes in pupils size (Di Stasi et al., 2015), different characteristics of blinking (May et al., 1990) and various characteristics of horizontal eye movements (May et al., 1990).

Analysed papers used eye-tracking to detect drowsiness, mental workload and cognitive distraction. The increased mental workload was associated with increased pupils' size, and a decrease in the number of blinks, spontaneous eye-movements and blinks duration (May et al., 1990; Palinko et al., 2010). Cognitive distraction was detected with PERCLOS algorithm (Di Stasi et al., 2015; Rodríguez-Ibáñez et al., 2011), decreased pupils size (Kristjansson et al., 2009) and reduced the speed of micro-saccades (Di Stasi et al., 2015). PERCLOS was defined as a proportion of time when eyes were closed over a certain period (Abe et al., 2014). However, PERCLOS was reported effective in the detection of distraction; it was mainly a tool to detect drowsiness (Abe et al., 2014; Brookhuis & de Waard, 2010). Drowsiness could also be detected with an increased blink duration, delayed lid opening, decreased lid closure speed, increased number and speed of saccades, increased number of off-road fixation (Schleicher et al., 2008; Wang et al., 2017). There were some differences in results related to drowsiness and pupils' size. Some authors reported an increase (Wang et al., 2017) and some a decrease (Oken et al., 2006).

3.3.4 ELECTROCARDIOGRAPHY

ECG is a measure of the electrical activity of the heart (Saritha et al., 2008).

Eighteen identified papers used some properties of heart functioning as a state indicator. The states identified with ECG were mental workload, drowsiness, fatigue, behavioural distraction, stress and anger. They were either identified using heart rate (Averty et al., 2002), inter-beat interval (Veltman & Gaillard, 1996), heart rate variability (HRV) (Wilson, 2002), or machine learning algorithms with multiple heart-related features (Sahayadhas et al., 2015). HRV is a measure of a natural, physiological variation in the heart rate (Roscoe, 1992). Drowsiness and fatigue were presented to correlate with a decreased (Maglione, Borghini, Arico, et al., 2014; Ogorevc et al., 2011), while stress with an increased heart rate (Schreinicke et al., 1990). Both anger and distraction were identified with machine learning algorithms using multiple features (Minhad et al., 2017; Sahayadhas et al., 2015). The increased mental workload was found to be correlated with increased heart rate (Averty et al., 2002), decreased heart rate variability (Roscoe, 1992) and decreased inter-beat interval (Veltman & Gaillard, 1996).

3.3.5 ELECTROOCULOGRAPHY

EOG is a measure of the ocular behaviours through the resting potential of the retina (Siddiqui & Shaikh, 2013). Similarly to eye-tracking, it provides data about blinks and horizontal eye movements, but it does not show changes in pupil size.

Sixteen of the identified papers measured state with EOG. The measured states were drowsiness (Borghini et al., 2014), fatigue (Lal & Craig, 2001), sleep (Oken et al., 2006) and mental workload (Richter et al., 1998).

The increased mental workload was associated with decreased blinking rate and blinking duration (Borghini et al., 2014). Drowsiness was reported to be correlated with the decreased saccadic eye-movements, increased slow eye-movements, increased blinking duration, delayed lid opening, and decreased lid closure (Borghini et al., 2014; Schleicher et al., 2008). It was also identified with PERCLOS (Papadelis et al., 2007). Two papers reported an increase in the blinking rate due to drowsiness (Borghini et al., 2014; Papadelis et al., 2007), but one reported a decrease (Minhad et al., 2017). Fatigue was reported to be associated with the increased blinking speed, the disappearance of saccadic eye movements, and an increase of PERCLOS (Lal & Craig, 2002; Rodríguez-Ibáñez et al., 2011). One paper reported a rise in the blinking rate (Stern et al., 1994), while one a decrease (Morris & Miller, 1996). Sleep was identified with slow eye movements (Oken et al., 2006).

Most of the results were consistent with eye-tracking data except for the reported decrease of saccades number that contradicted the finding of Wang et al. (2017); however, saccades are very fast movements and their detection might depend on the sampling rate of the device.

3.4 CONCLUSIONS

This chapter reviewed available methods of driver state monitoring or of psychophysiological detection of the states previously identified as risky for driving. It was a crucial step for further review of the circadian influence on the driver state monitoring (see Chapter 6) and for the selection of measures for the experiment (see Chapter 7). It was also a basis for the further selection of the tools for the experiment. The formal process of how the experimental tools were selected was described in detail in chapter 7. EEG and eye-tracking were identified as frequently used methods, and later in Chapter 7, they were selected for the experiment due to their high usability. However, this thesis did not describe the results from them. Eye-tracking was excluded at the phase of the pilot study. It created a high computational load that disabled other measurements, as well as participants, experienced high discomfort due to the eye-tracking headset and were not able to participate in the full experiment. Electroencephalography was recorded during the experiment, but the analysis of the brain signals was not included in this work due to the high temporal load of the pre-processing, technical problems and lower data quality in comparison to other recorded signals.

4. CIRCADIAN EFFECT ON MANUAL DRIVING- REVIEW

4.1 INTRODUCTION

A number of psychophysiological states have been reported as potentially jeopardizing driving safety (see chapter 2). The most significant were, sleep, sleep inertia, drowsiness/sleepiness, fatigue, cognitive distraction, behavioural distraction, suboptimal mental workload and insufficient situation awareness. Sleep inertia, drowsiness/sleepiness and fatigue are closely related to sleep, and the driver might fall asleep as a result of these states (Higgins et al., 2017; May & Baldwin, 2009). Sleep undergoes circadian rhythmicity; hence, it is likely that driving risk factors also change over the 24-hour cycle. Also, the vast majority of the risk factors identified have a large physiological component. Therefore, understanding the physiology of the driver has the potential to provide insight into the mechanisms that underpin performance that we can define as the ‘state' of the driver. A critical task is defining the driver states that impact upon the capacity of the driver to optimally perform driving tasks, automation monitoring and in particular the transition between the two. A critical governor of an individual's physiology is the daily circadian cycle, promoting sleep alertness at specific points in time most frequently synchronised with day and night (Dijk et al., 1992; Van Dongen & Dinges, 2000). It could have profound implications on the ‘state' of the driver at different times of day but is rarely considered when examining automated driving. This chapter aimed to review the existing knowledge of the circadian rhythmicity in manual driving. It provided a base for reasoning about circadian rhythmicity in semi-automated driving.

4.2 METHODOLOGY OF THE LITERATURE SEARCH

The databases used were: Scopus, Web of Science, Google Scholar and DelphiS. The initial search terms used in the databases were the circadian effect on driving; car OR vehicle OR automotive. Those terms were purposely selected to be broad and create a good basis for further iterative research in references and terms suggested by the initially identified literature as recommended in grounded theory (Heath & Cowley, 2004). After the primary search, key references of the relevant publications were analysed to identify additional positions. The key references analysis was used iteratively. The initial number of results was, 32 in Web of Science, 28 600 in Google Scholar, 127 in DelphiS, and 83 in Scopus.

Inclusion Criteria: only full access articles in the English language were included in the criteria. Only 500 first results of the search in the database results were included. The included materials concerned circadian effect on driving performance.

Exclusion Criteria: The papers about the circadian effect on other kinds of operations than driving were excluded. Also, papers that analysed potential interventions in night-driving were not further evaluated.

As a result, 23 papers were included in the analysis.

4.3 RESULTS

Circadian variations in bodily functions depend on a variety of regulating mechanisms. Melatonin plays the role of an internal while the light of an external controlling factor of those rhythms (Blatter & Cajochen, 2007; Dijk et al., 1992). Studies have repeatedly reported that many cognitive and psychomotor functions decreased in the circadian night when the melatonin level is the highest. There was also a smaller cognitive decrease during the so-called ‘mid-afternoon dip'. These periods of times were also associated with the increased subjective feeling of being tired (Dijk et al., 1992; Van Dongen & Dinges, 2000). This chapter summarised papers addressing the influence of the circadian rhythms on manual driving.

That was not a trivial question to answer because of the confounding effect of sleep deprivation. Frequently, the activity observed during the circadian night followed a long period of no sleep; hence its quality might be influenced both by circadian rhythmicity and sleep deprivation (Dijk et al., 1992). Most of the studies analysed did not disentangle those two effects, leaving an open question about the proportions of circadian and sleep deprivation effects on the observed variables.

The approach taken in the studies fell into three basic types; an analysis of the police reports comparing day and night accidents, driving simulator experiments, and real-life driving experiments.

Analysis of the police reports and real-life driving experiments did not allow unambiguous identification of the causal factors. Drivers who drove at night could be sleep deprived more often than drivers who drove in the day. Therefore, it is was possible to evaluate what was the contribution of the circadian effect in the changes in their performance.

Experiments conducted in the driving simulator might use different experimental designs. Only two procedures allow for untangling the circadian effect from the sleep deprivation effect. Forced desynchrony protocol uses an isolated laboratory environment without time cues to change sleep/wake rhythm. The time for sleep and wake is increased in the way that participants who live in the laboratory respectively sleep in every part of the circadian cycle without sleep deprivation (Kosmadopoulos et al., 2017; Matthews, Ferguson, Zhou, Kosmadopoulos, et al., 2012; Matthews, Ferguson, Zhou, Sargent, et al., 2012). The alternative option is testing highly sleep-deprived participants in different circadian phases (Williamson & Friswell, 2008). To address this problem, the methods of the included studies were listed in Table 4.1.

As a result of the literature review, twenty-three papers that evaluated the effect of circadian rhythm or melatonin level on the different aspects of driving were included in further analysis. The most frequently described effect was a decrease of some functions or increase of negative influences during the circadian night and some particular times during the circadian day. Table 4.1 listed periods when respective risky states of the driver were experienced with higher intensity or were more likely to occur. There was a distinction in the table between Sleep and Sleepiness causes of the accident risk, as these were terms that have been used in the papers. They were mostly not defined; however, the literature defined sleepiness as a state in between sleep and wake that is associated with reduced performance and attention (Jackson et al., 2016).

Table 4.1: List of the driver's states that were observed to be affected by circadian rhythmicity, the reference to the relevant paper, and a brief description of the methodology of the study.

Risk factor	**Time of the day when increased**	**References**	**Methodology**
Distraction	2 a.m. – 7 a.m., 2 p.m. – 5 p.m.	(Mitler et al., 1988)	Analysis of the police and hospital reports
Sleep	2 a.m.- 6 a.m.; 2 p.m.- 4 p.m.	(Pack et al., 1995)	Statistical analysis of the police reports
	2 a.m. – 7 a.m., 2 p.m. – 5 p.m.	(Mitler et al., 1988)	Analysis of the police and hospital reports
	0-2:59 a.m. with a peak at 2 a.m., 6 a.m., 4 p.m.	(Horne & Reyner, 1995)	Analysis of police statistics and spot interviews
Sleepiness	11 p.m. – 1 a.m.	(Otmani et al., 2005)	Driving simulator experiment
	Night	(Lowden et al., 2009)	Driving simulator experiment among young and older drivers. Young drivers are more susceptible to night sleepiness.

Risk factor	Time of the day when increased	References	Methodology
	Night, with a peak around 3.30 a.m.	(Sandberg et al., 2011)	Real-life driving experiment
	Night, increasing with increasing time of wakefulness	(Akerstedt & Folkard, 1995)	Real-life driving experiment
	Night	(Akerstedt et al., 2013)	Real-life driving experiment
	0 a.m.- 2 a.m., 3 a.m.- 5a.m.	(Sahayadhas et al., 2013)	Real-life driving experiment
	Night	(Gillberg et al., 1996)	Driving simulator experiment
Fatigue	Night	(Williamson & Friswell, 2008)	Sleep deprivation in different circadian phases experiment
Impairing effect of alcohol consumption	Night and early morning	(Garbarino et al., 2016)	Testing the level of alcohol in breath in the drivers involved in the car accident. Analysis of circadian effect and comparison to the control group.
Generally the low driving quality	Night, with a peak 4 a.m.-6 a.m.	(Akerstedt & Kecklund, 2001)	Statistical analysis of the accidents data.
	2 a.m. – 7 a.m., 2 p.m. – 5 p.m.	(Mitler et al., 1988)	Analysis of the police and hospital reports
	2 a.m.- 5 a.m.	(Chipman & Jin, 2009)	Analysis of the police reports
	Early morning, with a peak at 4 a.m.	(Akerstedt et al., 2001)	Analysis of the police reports
	Night	(Williams, 1985)	Analysis of the accidents data from the Bureau of the Census
	Night	(Matthews, Ferguson, Zhou, Kosmadopoulos, et al., 2012)	Forced desynchrony with sleep restriction protocol
	Night	(Matthews, Ferguson, Zhou, Sargent, et al., 2012)	Forced desynchrony with sleep restriction protocol
	8 a.m.- 10 a.m. after a night shift	(Akerstedt et al., 2005)	Driving simulator experiment
	Night	(Gillberg et al., 1996)	Driving simulator experiment

Driving performance generally improved throughout the day (Lenné et al., 1998) with a decrease between 2 p.m. and 5 p.m. (so-called 'mid-afternoon' dip) (Mitler et al., 1988) and then a more substantial reduction at night (with the lowest scores between 2 a.m. and 5 a.m. and accidents risk peak at 4 a.m.) (Akerstedt & Kecklund, 2001). Figure 4.1 presented an example of a graphical approximation of the driving performance in different circadian phases. It is important to note that the methodology of this work did not allow to draw a precise curve of performance or to establish local extrema, a slope of the curve, or the inflexion points.

Even though there were fewer vehicle collisions during the night, the risk of getting into an accident was significantly higher at this time. Such accidents risk could not be explained solely by reduced visibility (Akerstedt & Kecklund, 2001; Chipman & Jin, 2009).

The groups that were at the highest risk were young males (Akerstedt & Kecklund, 2001) and young people in general (Otmani et al., 2005).

In terms of risky driver states, sleepiness was significantly increased at night (Akerstedt et al., 2013), especially in young drivers (Lowden et al., 2009). The peak of the night-time sleepiness was observed around 3.30 a.m. (Sandberg et al., 2011). Akerstedt and Folkard (1995) reported that sleepiness was gradually increasing at night with growing wakefulness time and reached its peak at the end of the night shift. Exposure to the bright light, which decreased the level of melatonin, also reduced the amount of night-time fatal accidents in driving simulators (Leger et al., 2009; Weisgerber et al., 2017). Driving research showed that certain types of accidents had distinctive frequency changes during the twenty-four hours cycles. Most of the sleep-related crashes occurred between 2 a.m. and 6 a.m., and then from 2 p.m. to 4 p.m. (May & Baldwin, 2009). Circadian rhythm interacted with the

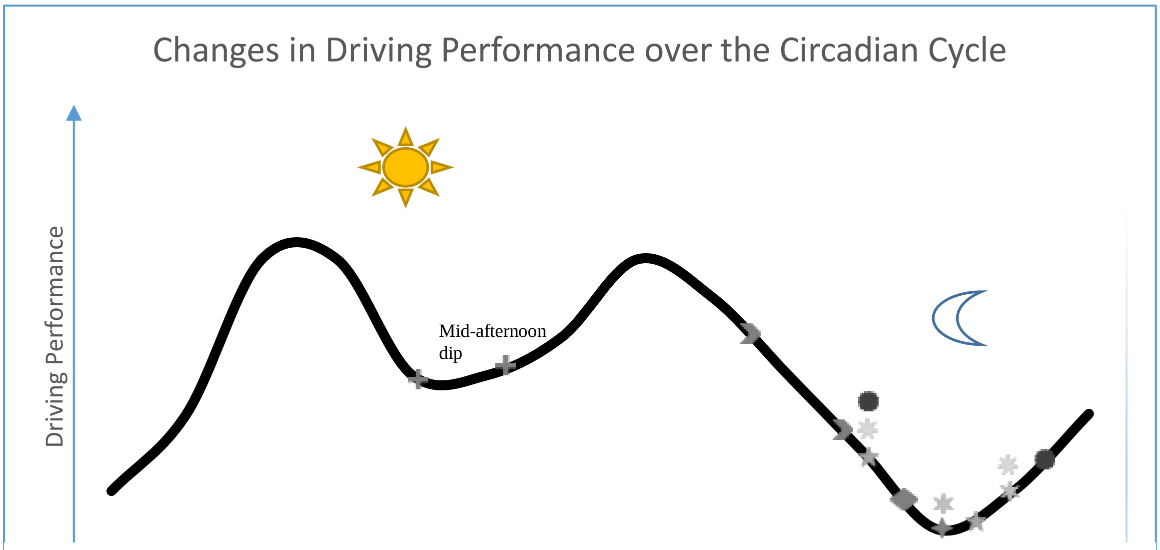

Figure 4.1: A sample curve approximating the changes in driving performance over 24-hours. The graph shows the worst driving performance period between 3.30 and 5. ✚- 2p.m.-5p.m. mid-afternoon dip, peak of accidents, higher probability of falling asleep, peak of distraction-related accidents (Horne & Reyner, 1995; Mitler et al., 1988; Pack et al., 1995, p. 19) ➤- 11 p.m.- 1 a.m. peak of sleepiness (Otmani et al., 2005) ★- 2 a.m.- 5 a.m. peak of accidents (Mitler et al., 1988) ✳- 2 a.m.- 6 a.m. the highest risk of falling asleep (Horne & Reyner, 1995; Pack et al., 1995) ●- 2 a.m.- 7 a.m. the highest risk of falling asleep (Mitler et al., 1988) ◆- 3.30 peak of sleepiness (Sandberg et al., 2011) ✦- 4 a.m. peak of accidents (Akerstedt et al., 2001) ✳- 4 a.m.- 6 a.m. peak of accidents (Akerstedt & Kecklund, 2001).

effects of alcohol consumption, which made collisions more likely to when drinking and driving at night rather than during the day (Garbarino et al., 2016).

There are few studies that disentangled the circadian effect from sleep deprivation effect on driving. Matthews, Ferguson, Zhou, Kosmadopoulos et al. (2012), Matthews, Ferguson, Zhou, Sargent, et al. (2012) and Kosmadopoulos et al. (2017) analysed the data based on the forced desynchrony experiment with sleep restriction protocol. They found that the circadian phase had a significant effect on driving performance only when the driver has been awake for a long time or sleep-deprived. It means that drivers had more collisions if they drove at night, but only if they were awake for a long time prior to the driving or they did not have a sufficient amount of sleep. Williamson et al. (2008) found a similar effect in their study based on sleep deprivation in different circadian phases.

4.4 CONCLUSIONS

This chapter provided evidence that circadian rhythmicity affects manual driving. However, there is no data on the circadian effect on semi-automated driving. Whilst, it might seem that many tasks in semi-automated driving are similar to manual driving and hence could be affected by circadian rhythmicity, such a statement requires further investigation. Chapters 5 and 6 provided literature reviews on the topic of the circadian effect on semi-automated driving and driver state monitoring. Chapters 8, 9, and 10 presented experimental results related to the role of the circadian rhythmicity in the time-course of semi-automated driving (chapter 8), driver state monitoring (chapter 10), and general circadian differences in driving performance and driver physiology in semi-automated driving.

5. CIRCADIAN EFFECT ON SEMI-AUTOMATED DRIVING- REVIEW

5.1 MODELS OF DRIVER IN MANUAL DRIVING MODE

A variety of studies attempted to introduce a human driver model. Many models were related to the particular environment, like intersection (Liu & Ozguner, 2007) or specific cognitive context, like perceived risk (Liu & Ozguner, 2007). The attempts to create a universal model resulted in the cognitive types of models that focused on the cognitive processes that underlie the behaviours of the driver. One of the widely used models was COSMODRIVE, presented in figure 5.1, developed by the French Institute of Transportation Research and Safety INRETS. It was based on the framework from cognitive psychology and ergonomics (Delorme & Song, 2001). It divided the cognitive functions of the driver into different modules: strategic, tactical, operational, emergency management, management and control, perception, and execution.

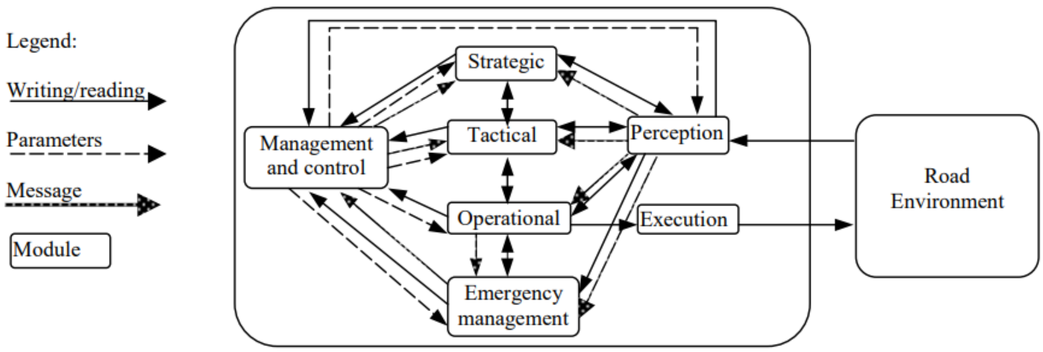

Figure 5.1: The general architecture of the COSMODRIVE from (Delorme & Song, 2001).

Each of the modules was a complex and experimentally motivated construct, as shown in figure 5.2 on the example of the Perception and Tactical Modules.

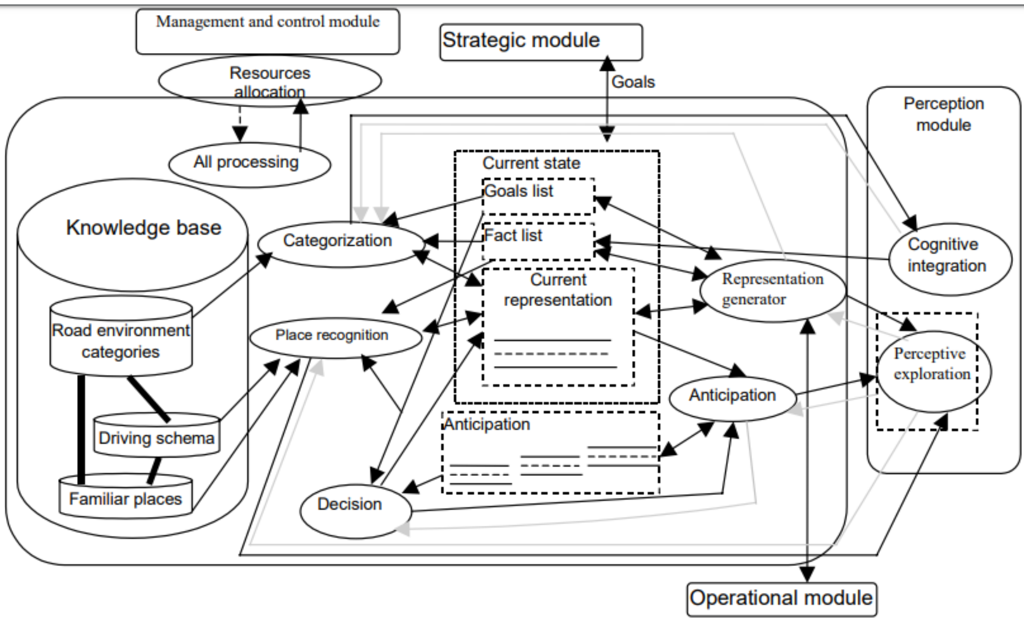

Figure 5.2: Schematic Diagram of the Perception and Tactical Modules in COSMODRIVE. Model from (Delorme & Song, 2001).

Some researchers attempted to simplify the COSMODRIVE to achieve a better practical impact. An example could be a Human Driver Model for SmartAHS that mainly concentrated on the Perception and Tactical Modules in the modelling of the driver cognition (Delorme & Song, 2001). Another widely deployed cognitive model of the driver was the Driver Behaviour Model that is used by the human factors team from TNO. As shown in figure 5.3, it stressed the importance of factors like visual attention, workload, comfort, and acceptance (Keith et al., 2005).

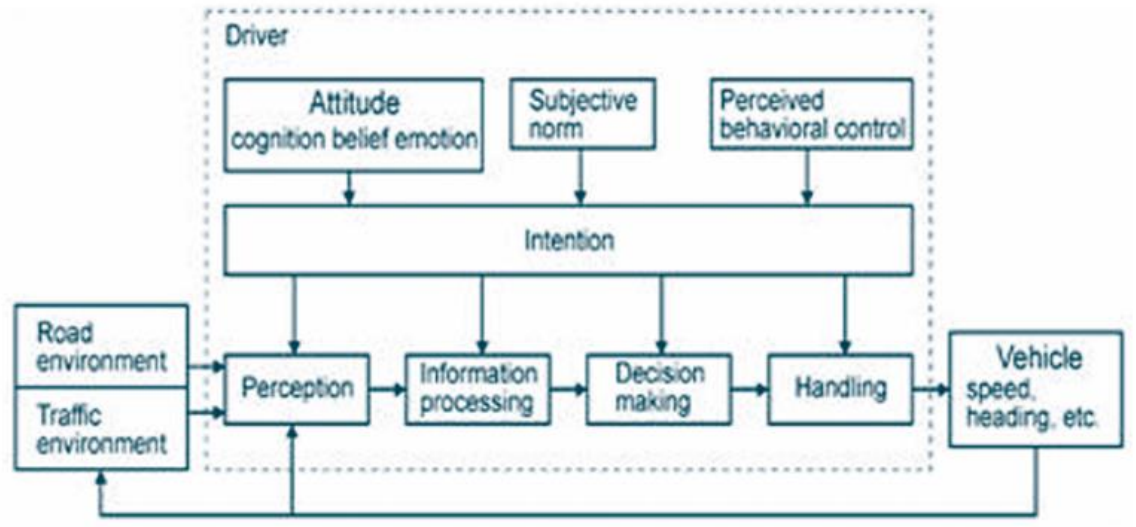

Figure 5.3: TNO driver behaviour model from (Keith et al., 2005).

These examples did not cover all the models of the driver; however, they are widely approved by road safety and human factors research institutes. They all created a picture of a complex cognitive process involving functions like attention, mental model, workload or perception (Delorme & Song, 2001; Keith et al., 2005; Liu & Ozguner, 2007).

5.2 MODELS OF THE DRIVER IN THE AUTOMATION

Understanding the driver role and state in semi-automated driving is essential for addressing the challenges and designing safer systems (Endsley, 2019). There are a variety of models describing human-automation interactions; however, not many models concentrate directly on semi-automated driving (Heikoop, de Winter, van Arem, & Stanton, 2016).

Stanton and Young (2000) proposed a model of the driver when using vehicle automation. They utilized existing literature to select the most significant psychological factors that can influence driver and driving in a semi-automated vehicle. They modelled the direction of the interactions between the selected factors. As shown in figure 5.4, the factors identified as the most significant were: task demands, stress, feedback, the locus of control, situation awareness, mental workload, trust, and mental model. Situation awareness was affected by the trust of automation, mental model, stress and mental workload. Stress and mental workload both influenced each other. Stress was also affected by task demands and locus of control, while the mental workload was affected by feedback (Stanton & Young, 2000).

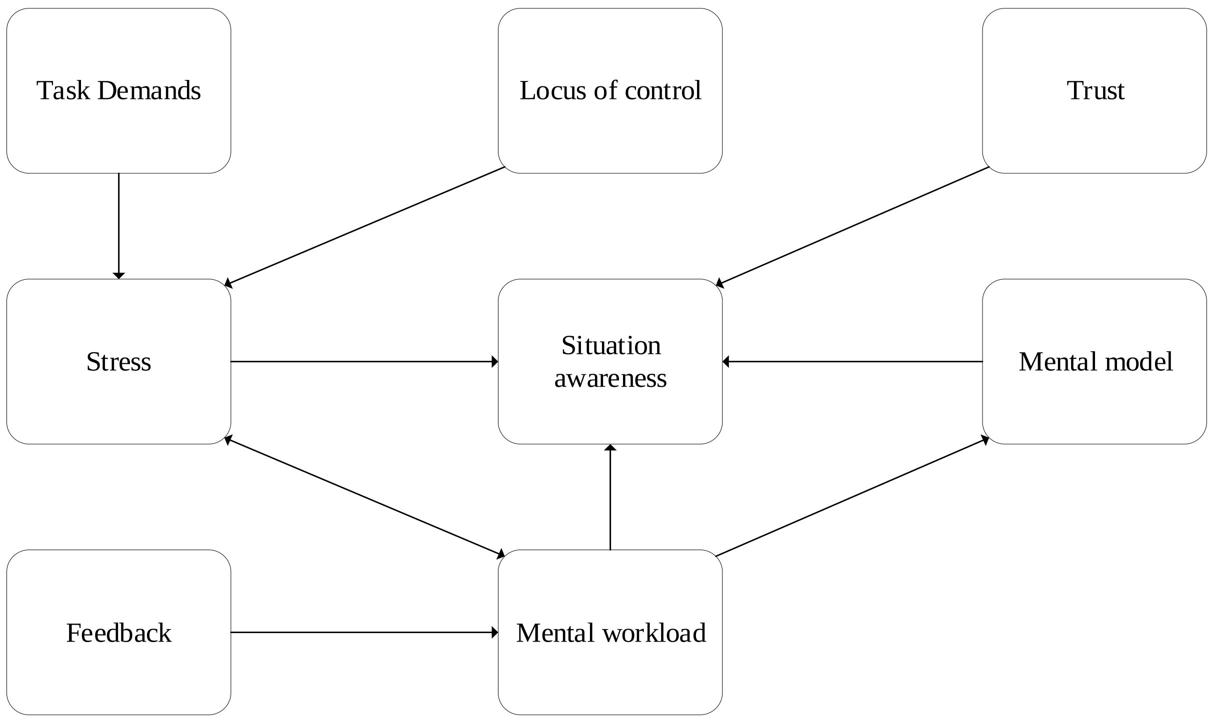

Figure 5.4: Graphical representation of driver in automation psychological model by Stanton and Young (2000).

HASO model depicted interactions between the automated system and operator, highlighting the most important states and factors in the system design and the environment. As shown in figure 5.5, the critical operator states in the HASO model were situation awareness, mental model, engagement, workload, attention allocation, and trust of automation. Together with different parts of the system design, they led to an optimal or suboptimal automation oversight and interaction performance (Endsley, 2016).

Unlike the model by Stanton and Young, it did not only concentrate on the driver in automation but on any type of operator of the automated or semi-automated system including the design of the system and external parts of the environment. Both models included situation awareness, workload, mental model, and trust as essential states of the operator. However, the model by Stanton and Young (2000) also included task demands, stress, feedback, and locus of control, while the HASO model included engagement and attention allocation.

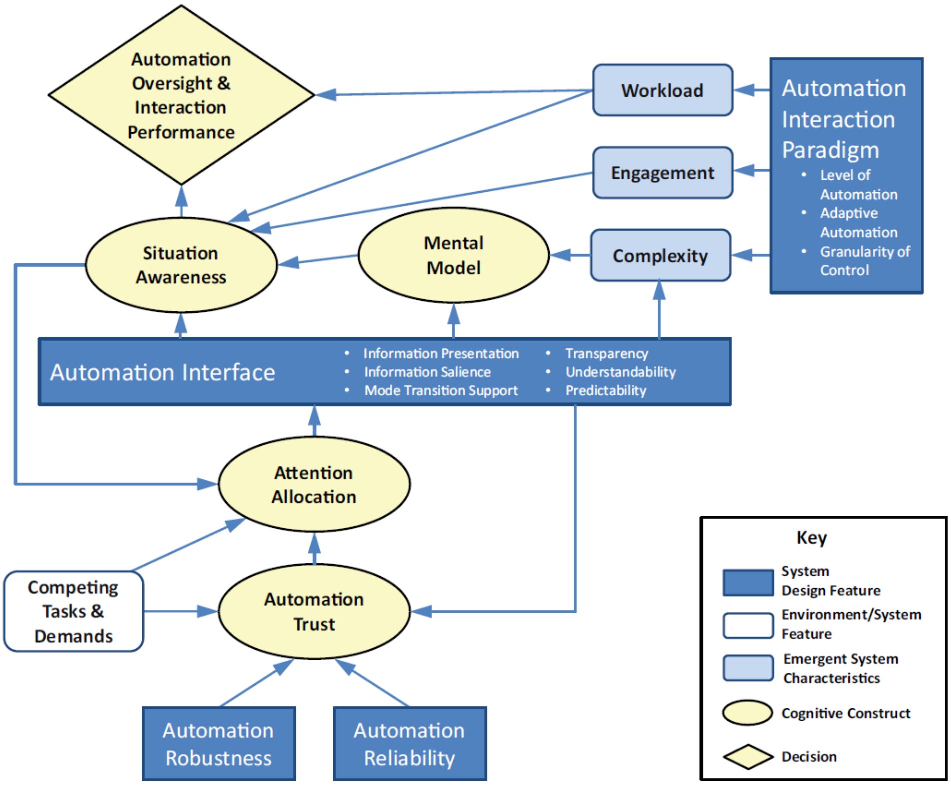

Figure 5.5: Graphical representation of the HASO model (Endsley, 2016)

Heikoop et al. (2016) updated the model by Stanton and Young (2000) and complemented it with the causal links. As a result, the Consensus Model was proposed, as shown in figure 5.6. The consensus Model did not include locus of control as a factor, but it added fatigue and attention as significant factors. Attention allocation was also a factor included in the HASO model (Endsley, 2016). The causal links were based on a thorough literature review (Heikoop et al., 2016).

This chapter used the Consensus Model as a base for further investigation into semi-automated driving. One of the reasons was that it is a comprehensive and contemporary model created as a result of an immense literature review. Unlike the HASO model, it is specific for a driving domain and entirely concentrated on the state of the driver. As also, it included fatigue as an important factor (Endsley, 2016; Heikoop et al., 2016), which is highly influenced by a circadian phase (Lowden et al., 2009; Otmani et al., 2005) and thus important for the topic. However, this model did not include all the factors involved in semi-automated driving, and more research might be necessary. Nevertheless, the aim of this chapter was to add circadian context to the Consensus Model without applying any changes to the model or the factors included in it.

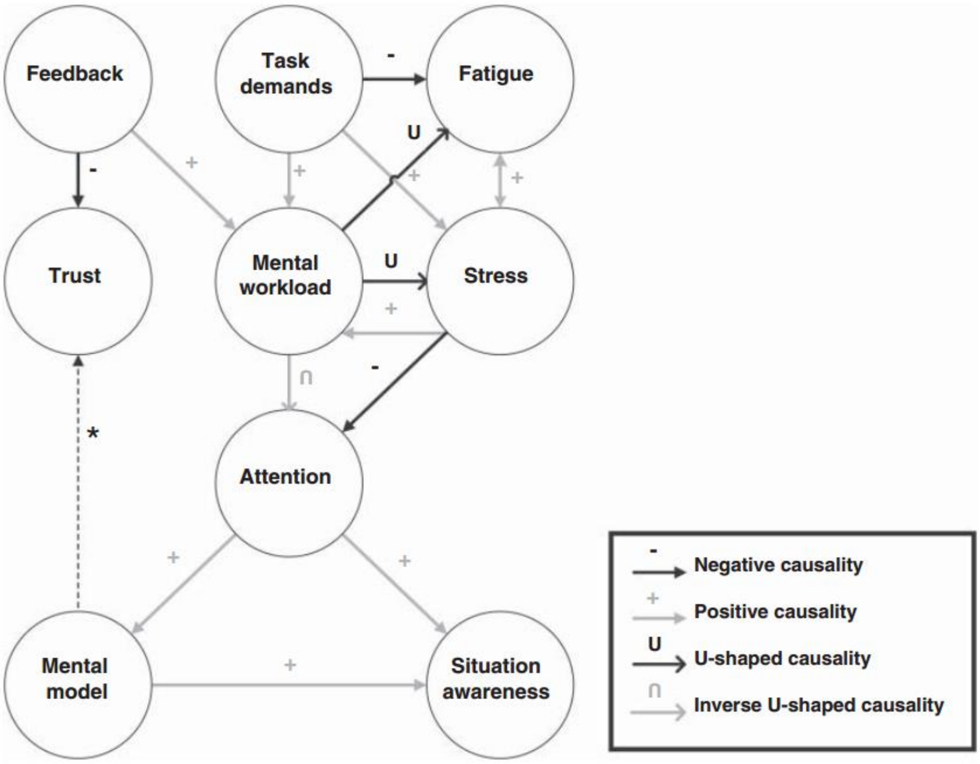

Figure 5.6: Graphical representation of the consensus model (Heikoop et al., 2016).

5.3 METHODS OF THE LITERATURE REVIEW

In their model, Heikoop et al. (2016) identified nine factors significant for the driver in automation, feedback, trust, mental model, task demands, mental workload, attention, fatigue, stress, and situation awareness. Each of these factors was used as a search term in Scopus and Web of Science databases. The search terms were a name of one of the factors (for example task demands) AND circadian effect AND driv* OR car OR vehicle. The first three hundred results were evaluated based on abstracts from the perspective of inclusion and exclusion criteria.

Inclusion criteria: papers included for evaluation were peer-reviewed publications in English with full text available. Both experimental articles and literature reviews were included. Papers had to evaluate or analyse the effect of one or more of the factors from the Consensus Model on driving performance in the circadian context.

Exclusion criteria: papers that only evaluated the circadian effect on driving performance without one of the factors from the model were excluded from the analysis. Similarly, papers that only evaluated the circadian influence on the factor from the model, but not in the driving environment were excluded. Publications that analysed different interventions to increase driving performance or used clinical groups were also excluded from the analysis.

As a result, only nine papers were identified as meeting the criteria. Because of this, a second search was performed. The second search extended the area of interest to pilots and operators of other types of vehicles or machines. Search engines again were Scopus and Web of Science. Search terms used in the second search were, a name of one of the factors AND circadian effect AND pilot OR operat* OR air* OR aviati*. The first three hundred results were evaluated based on abstracts from the perspective of inclusion and exclusion criteria.

The inclusion and exclusion criteria were the same as for the driving papers.

As a result of the two searches, fourteen (one paper included two factors) papers were selected for evaluation. The table presented the number of papers found and selected in each of the searches.

Table 5.1: Number of results given by search engines and the number of papers selected in each of the searches. The number of papers selected for further evaluation is presented in brackets. One of the paper included two factors; hence, the table included fifteen results in sum.

Factor from the Consensus Model	**Web of Science: circadian effect on factors in driving**	**Scopus: circadian effect on factors in driving**	**Web of Science: circadian effect on factors in different machine operators**	**Scopus: circadian effect on factors in different machine operators**
Fatigue	84 (2)	4667 (4)	156 (2)	7455 (2)
Attention	53 (1)	4731 (1)	68 (0)	6230 (1)
Mental Workload	0 (0)	582 (1)	4 (0)	781 (0)
Feedback	105 (0)	1652 (1)	45 (0)	5047 (0)
Trust	2 (0)	79 (0)	32 (0)	163 (0)
Mental Model	6 (0)	641 (0)	10 (0)	1043 (0)
Stress	149 (0)	17152 (0)	208 (0)	22010 (0)
Situation Awareness	3 (0)	238 (0)	2 (0)	302 (0)

Factor from the Consensus Model	Web of Science: circadian effect on factors in driving	Scopus: circadian effect on factors in driving	Web of Science: circadian effect on factors in different machine operators	Scopus: circadian effect on factors in different machine operators
Task Demands	5 (0)	158 (0)	13 (0)	1026 (0)

5.4 RESULTS

To analyse the circadian effect on semi-automated driving each factor from the Consensus Model (Heikoop et al., 2016) was investigated in the context of circadian effect and driving performance. The factors affected by the circadian phase directly or indirectly were discussed below.

5.4.1 FATIGUE

Fatigue is a state caused by the prolonged performance of one task. It can comprise of feeling tired, sleepy, need to stop the task, and decreased cognitive or muscular performance (Johns, 2000).

Ten papers related to fatigue were selected for further analysis. They either evaluated the association between circadian rhythms and fatigue in the driving domain, piloting, or machine inspection tasks. The results presented in the papers were highly consistent.

Fatigue significantly increased at night in drivers (Lowden et al., 2009; Otmani et al., 2005; Phipps-Nelson et al., 2011), pilots (Caldwell, 2005; Gander et al., 2015; van den Berg et al., 2016) and machine controllers (Drury et al., 2006). One study also observed a significant fatigue increase in drivers during the mid-afternoon dip period (Zhang et al., 2017). This trend was visible in performance, subjective fatigue reports and physiological measures (Caldwell, 2005). Studies that allowed dissociation between circadian effect and sleep deprivation found that fatigue increased at night, but only in interaction with higher sleep deprivation (Matthews, Ferguson, Zhou, Sargent, et al., 2012; Williamson & Friswell, 2008). Fatigue also increased when the mental workload was higher than optimal, and there is evidence that mental workload is higher at night (Otmani et al., 2005).

5.4.2 ATTENTION

According to Wickens et al. (2015) ‘Attention may be described by the metaphor of a searchlight (Wachtel, 1967). Two properties of the searchlight are relevant: its breadth and direction. The beam’s breadth can be subdivided into two components: that which we want to process (focused attention), and that which we must process but do not want to (divided attention). The direction of the searchlight—how it knows when, what, and where in the environment to illuminate—describe the properties of selective attention.' (Wickens et al., 2015, p. 70).

Three papers were identified in the area of the effect of the circadian phase on distraction and driving. One of them analysed the police reports about distraction-related accidents. The second was a detailed analysis of distraction from the real-life 100 cars study (Klauer et al., 2006). The third paper analysed sustained attention in the shift workers. Driver's distraction was found to be the highest between 2 a.m. and 7 a.m. with a smaller effect during the mid-afternoon dip between 2 p.m. and 5 p.m. (Mitler et al., 1988). Also, Klauer et al. (2006) found that distraction-related to secondary tasks had a higher impact on driving at night. Shift workers had a much worse cognitive performance during the night shift. A significantly affected function was sustained attention (Chellappa et al., 2019). Sustained attention is critical in semi-automated driving, because of the long periods when a driver might be required to stay vigilant without the subsequent performance of the manual driving tasks (Kyriakidis et al., 2019; Warm et al., 2008; Young & Stanton, 2002). The night could then be even a more risky period for semi-automated driving than for manual driving (Otmani et al., 2005). Interaction with Mental Workload depicted in the Consensus Model (Heikoop et al., 2016) also suggests a decrease of attention at night, due to the increase of mental workload (Otmani et al., 2005).

5.4.3 MENTAL WORKLOAD

One paper was selected in the area of circadian effect on mental workload and driving (Otmani et al., 2005). The mental workload of professional drivers was assessed after the day-time and night-time drives using the NASA-TLX. The mental workload was rated by the drivers as significantly higher after the night-time drive. The study design did not allow a dissociation between the circadian effect and sleep deprivation (Otmani et al., 2005). However, studies on the nuclear plant supervisors have provided evidence that supervising strategies significantly differed between the day-time and the night-time shifts (Andorre & Quéinnec, 1998; Andorre-Gruet et al., 1998). One of the possible explanations is a compensating strategy for the increased mental workload during the lower circadian periods, as participants declared being more tired during the night. In the case of mental workload increase, it is not straight-forward to predict its effect on driving performance. The highest performance occurs when the mental workload is on the optimal level (Young & Stanton, 2005), thus increased mental workload might have a positive or negative effect on performance.

5.4.4 FEEDBACK

The definition of feedback was often dimensionalised as it can describe a variety of information given to the user (Heikoop et al., 2016).

One paper that evaluated the feedback and circadian rhythmicity in driving was identified (Aidman et al., 2015). Feedback is a broad term and its effect on driving might differ depending on the type, quantity and form of the feedback. The selected paper analysed the effect of drowsiness feedback on performance in military drivers. Feedback about drowsiness level decreased drowsiness and increased driving performance. However, the improvement in performance due to the feedback was only observed during the day. The effect of feedback gradually reduced in the evening. Unfortunately, the authors did not have sufficient night data to judge the night-time effectiveness of the feedback (Aidman et al., 2015). Thus it is difficult to draw a general conclusion about how effective feedback is in different circadian phases. However, considering the general decline of cognitive performance in the night (Dijk et al., 1992) it could be hypothesized that cognitive processing of feedback is less effective in the night, but this requires empirical investigations.

5.4.5 TRUST

Trust was defined in a variety of ways, mostly including a state or approach characterised by vulnerability (Heikoop et al., 2016). Trust was repetitively documented as an important factor for automation (Endsley, 2016). Interaction with the automated system might be associated with two types of trust-related risks. Too high level of trust might lead to complacency and decreased situation awareness (Bailey & Scerbo, 2007; Endsley, 2016). At the same time, the too low trust might lead to disuse of the automation (Wiegmann et al., 2001). In fact, trust was proven to be one of the most important factors influencing the decision about the use or disuse of automation (Madhavan & Wiegmann, 2007).

It was a well-documented phenomenon that cognitive performance, as well as driving performance, dropped down during the night. Nonetheless, the literature that would combine the circadian effect with driving and different factors for semi-automated safety is quite scarce. Some of the factors from the Consensus Model were not analysed in such a combination; however, the model depicted the causality between the factors. To the knowledge of the author, trust was not yet analysed from the perspective of the circadian phase. Trust in the model by Heikoop et al. was negatively affected by the feedback and modulated by the mental model (Heikoop et al., 2016). Increased feedback decreased the level of trust for automation. Effectivity of feedback might be negatively affected by the circadian night (Aidman et al., 2015). It is not clear how would this, as a result, influence the trust.

5.4.6 MENTAL MODEL

An accurate mental model is essential to understand the environment and react adequately (Stanton & Young, 2000). The literature search did not allow to identify any papers on circadian effect on mental model and driving. However, the Consensus Model indicated that increased attention improved the use of the mental model (Heikoop et al., 2016). As attention decreases during the night (Dijk et al., 1992), it is likely that the retrieval and application of the mental model might be affected. The use and recovery of the mental model improves with improved attention (Heikoop et al., 2016). As attention decreases during the night (Mitler et al., 1988), the quality retrieval and applying of the mental model might also decrease.

5.4.7 STRESS

Stress was repetitively ill-defined; however, most of the definitions described it as a physiological and psychological reaction to the perceived danger (Heikoop et al., 2016; Matthews, 2002). No literature was identified in the area of a circadian effect on stress and driving. However, the consensus model depicted that fatigue as well as mental workload when higher than optimal, increased stress (Heikoop et al., 2016). Both fatigue and mental workload increased at the night (Dijk et al., 1992; Otmani et al., 2005). It is not easy to predict what would be an effect of increased mental workload on stress, as the relationship between them is a ‘U' shaped function.

5.4.8 SITUATION AWARENESS

Endsley (1996) defined situation awareness as ‘the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future'. She divided situation awareness into the three levels: level 1 based on the perception of the situation, level 2 based on the correct interpretation of the perceived stimuli, and level 3 based on the effective prediction of the near future. The deficit in any of those three levels defined an insufficient situation awareness (Endsley, 1996). Other authors simply described situation awareness as ‘knowing what is going on' (Stanton & Salmon, 2009). There were also no publications identified on the topic of circadian effect on situation awareness and driving. It is possible that situation awareness decreases at night as a result of reduced attention.

5.5 CIRCADIAN EFFECT ON THE CONSENSUS MODEL

The literature search showed that some factors from the Consensus Model undergo circadian fluctuations in the driving environment. Fatigue increased at night and mid-afternoon dip, attention decreased at night and mid-afternoon dip, the mental workload increased at night, and the effectivity of the feedback dropped at night. Trust, stress, mental model and situation awareness might also change with circadian phases because the other factors from the model influence them. However, it is not always easy to predict the effect of the fluctuation. The increase of the mental workload might lead to increased stress if the baseline mental workload was in the optimal range. It is also not easy to predict changes in trust as a result of the decreased effectivity of feedback and degraded mental model. Table 5.2 summarized circadian effects on the factors in the Consensus Model. Each row contained one factor. Each column represented the circadian phase: day, mid-afternoon dip, and night. Numbers in the cells represent numbers of papers identified during the literature search that evidenced an increase of these factors over a particular phase (number on the left), or the decrease (number on the right). Numbers in italics represented a number of papers that identified the circadian effect on factors influencing this particular one. For example, the mental model row had *2/0* in the day column because two articles identified that attention increased during the day, while according to the Consensus Model attention positively influenced stress. Question marks in the Trust represented the difficulty in modelling the effects of feedback and mental model changes on trust. Question marks in the Stress road represent the difficulty in modelling changes in stress without knowledge about the baseline level of Mental Workload.

Table 5.2: Summary of the circadian fluctuations of the factors from the Consensus Model

	Circadian Day	**Mid-Afternoon Dip**	**Circadian Night**
Fatigue	↑0/10↓	↑1/0↓	↑10/0↓
Attention	↑2/0↓	↑0/1↓	↑0/3↓
Mental Workload	↑0/1↓	↑0/0↓	↑1/0↓
Feedback	↑1/0↓	↑0/0↓	↑0/1↓
Trust	?	?	?
Mental Model	↑*2/0*↓	↑*0/1*↓	↑*0/2*↓
Stress	?	?	?
Situation Awareness	↑*2/0*↓	↑*0/1*↓	↑*0/2*↓
Task Demands	↑0/0↓	↑0/0↓	↑0/0↓

It is important to note that the studies included in this literature review did not allow precise modelling of the circadian effect. The studies used different methods to measure changes in factors like fatigue or attention; therefore, it is not possible to assess the magnitude of the change. The only indicator of the magnitude could be the frequency of the literature descriptions.

Papers also either aggregated results from long periods, measured factors at only particular time-moments or did not even give precise time information but used terms like ‘night’ or ‘day’. As a result of this, available data did not allow modelling of changes, local extrema or exact inflexion points. Therefore, this chapter presented an example of an approximation of the circadian changes in the Consensus Model. The size of the changes, as well as local extrema, inflation points and slope, were beyond the scope of this chapter and require additional experimental research.

The Consensus Model is a one-period model that does not include circadian rhythms or other time-related fluctuations. The proposition presented in this chapter offers a multi-period Consensus Model that included a circadian effect on the factors (Aronson, 1986). The multi-period model was given for a basic period of 24 hours. The underlying graph is the Consensus Model from Heikoop et al. (2016). The interactions between the factors stayed the same; however, the level of the factors changed depending on the circadian phase. Figure 5.7 presented the multi-period Consensus Model over four time periods, T1- day (before the mid-afternoon dip), T2- mid-afternoon dip, T3- day (after mid-afternoon dip), and T4- evening and night. The time was represented on the horizontal axis.

The vertical axis represented the performance of the driver in automation. The higher the model was placed in relation to the vertical axis, the better the performance was. The factors that were highlighted in blue grating increased in a particular period, while factors depicted in solid red decreased. Arrows placed next to the factors also represented the increase or decrease with a number above representing a number of papers that described such an effect. This frequency indicator could also be treated as an initial attempt to model the magnitude of the effect. Some arrows within the model were highlighted in blue, indicating a positive effect of the factor on the other factors. The proposed model did not describe the full scope of the circadian rhythmicity within the Consensus Model, but an initial proposition of the most prominent diurnal fluctuations. The literature review extracted factors from the Circadian Model that were documented to change within the circadian cycle in the driving domain and the ones that might change due to the interactions with other factors. Only the factors directly reported to fluctuate within the circadian cycle were included in circadian modelling. The exception was the effect of circadian rhythmicity on Feedback that was not indicated in the model. The reason was the unclear shape of the interaction. Also, the feedback described in the Consensus Model was related to automation, while the type of feedback evaluated from the circadian literature was associated with the state of the driver. The effect of the mid-afternoon dip was included in the model, even though it was smaller and less documented than the effect of the circadian night. However, there was a reported increase in road accidents in the afternoon that cannot be fully explained by the changes in traffic. At the same time, the number of studies tackling that phenomenon is even smaller than the number of studies on night driving. Because of that, it is essential to include it in the model to avoid the complete omission of this problem.

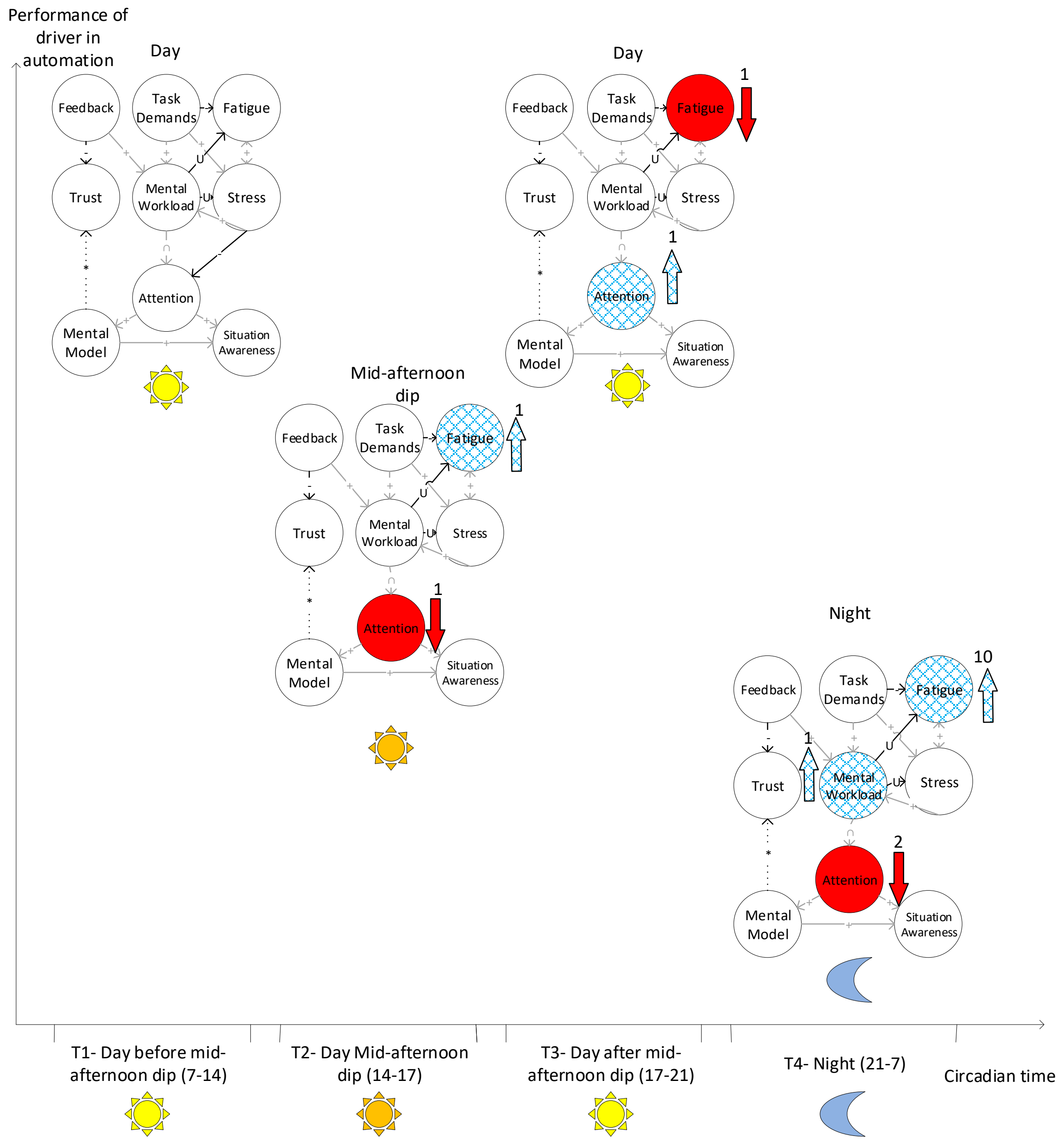

Figure 5.7: Multi-period Consensus Model representing circadian fluctuations of factors. The vertical axis represented the performance of the driver in automation. The horizontal axis represented circadian time, and it was divided into four periods: T1- the day before the mid-afternoon dip (7-14), T2- Mid-afternoon dip (14-17), T3- the day after mid-afternoon dip (17-21), and T4-night (21-7). Factors that increased in a particular phase were highlighted in the blue grating, and factors that decreased were highlighted in solid red. The arrows additionally indicated the direction of change. The number above the arrow represented the number of evaluated papers that described this process.

The multi-period Consensus Model proposed similar fluctuations of semi-automated driving performance to manual driving performance. Night and mid-afternoon dip might be characterised by a decreased performance due to the increased fatigue, decreased attention, decreased feedback effectivity and increased mental workload.

5.6 CONCLUSIONS

Humans are naturally diurnal animals and their cognitive performance decreases at night (Blatter & Cajochen, 2007; Dijk et al., 1992, 1997). Both statistical analyses of collisions and experimental studies provided evidence that night-time manual driving is significantly more dangerous than day time driving (Akerstedt & Kecklund, 2001; Horne & Reyner, 1995; Mitler et al., 1988). Surprisingly, the research related to driving and circadian rhythmicity is scarce. Moreover, the methodology of the majority of the studies did not allow dissociation between circadian effect and sleep deprivation. Many experiments did not model detailed changes in driving performance over hours but just differentiated between day and night.

Additionally, statistical analysis of the police reports did not enable an in-depth understanding of the collisions’ causality. To the knowledge of the author, there is no research investigating the circadian effect on semi-automated driving. Such negligence of the importance of the circadian effect is not specific for a driving domain. A similar phenomenon could be observed in pharmacology, where the circadian phase might highly modulate the effect of the drug, but the research in this area is scarce (Lemmer, 1995). This chapter proposed a multi-period, Consensus Model, of the circadian effect in driving automation. The literature review suggested an increased risk in automated vehicles at night and during the mid-afternoon dip due to the decreased attention, increased fatigue and increased mental workload. The multi-period Consensus Model proposed a theoretical background for the experimental studies in this domain. This chapter was also a call encouraging more work in this area. Negligence of this topic leads to a lower understanding of the phenomenon and decreased road safety. Future studies should provide more experimental evidence for the circadian changes in driver attention, fatigue and mental workload, and explore potential circadian fluctuations in other factors from the Consensus Model, for example, trust or feedback. It should also investigate driver performance in the semi-automated vehicles over different times of the day and night.

Additionally, the amount of studies dissociating between changes in driving performance due to the circadian influence and sleep deprivation is low. Because of that, it is still an open question if night driving riskiness is mainly caused by the circadian effect or by fatigue related to prolonged wakefulness. Answering such a question would give a better theoretical basis for actions aiming for an increase in driving safety. The understanding of how the circadian phase influences driver state, driving performance, ability to take-over and to cooperate with the automation could allow the better design of the automated systems. The artificial intelligence could support the driver differently during the night to compensate for the weakest points in their performance. Even though, some aspects of night driving, like reduced visibility, might present a challenge for the automated system, some other might be easier for the automation than for the human driver, for example, prolonged sustained attention and monitoring.

6. CIRCADIAN EFFECT ON DRIVER STATE MONITORING- REVIEW

6.1 INTRODUCTION

Driving performance and driving safety undergo a distinctive circadian fluctuation with night and mid-afternoon dip being characterized by worse performance and a larger accident risk (Akerstedt et al., 2001; Matthews, Ferguson, Zhou, Kosmadopoulos, et al., 2012; Mitler et al., 1988). Most of the physiological functions also undergo circadian variability (Frank et al., 1966), for example, the amount of cortisol decreases in the evening and at night and increases in the morning (Del Corral et al., 2016). As so, accuracy and interpretation of the driver state monitoring might be affected by the time of the day. Suppose the measured physiological functions had different baseline value depending on the time of the day (or night). In that case, the interpretation of its value should take a circadian phase into account. In a given situation, a physiological state associated with sleepiness during the day could be a normal baseline at night. Similarly, decreased driving performance at night should lead to more sensitive or careful monitoring systems in this circadian phase. Considering the danger of driving accidents, such conclusions are not merely theoretical considerations but might have severe and practical safety consequences.

This chapter reviewed the literature about circadian variations in physiological processes that underly measures used for driver state monitoring. It was undertaken to create the theoretical basis for experimental research in the area of circadian rhythms in driver state monitoring, as well as to encourage and justify the need for such research. The knowledge about the mechanisms and dynamics of physiological processes can ensure effectivity and accuracy of the monitoring systems that use them. There have been little experimental studies on driver state monitoring that has considered circadian rhythmicity. Considering the likely effects of circadian rhythms on driver physiology, the current lack of knowledge on the topic means that driving safety may be compromised. As such, this review does not only play a theoretical role but could also contribute to driving safety.

6.2 METHODS OF THE LITERATURE REVIEW

The primary strategy of this literature review was based on the grounded theory approach (Parnell et al., 2016; Rafferty et al., 2010). The work attempted to integrate knowledge about driver state monitoring, monitoring of the states that can jeopardize driving safety and circadian physiology. Because of that, there is no common framework that could be used as a method for a literature search. The grounded theory allows a broad, explorative approach and comprehensive perspective, which made it suitable for this research.

A search was meant to explore the circadian effect on the variety of measures used for driver state monitoring. It was conducted in Web of Science, Scopus and Pubmed. The selection of the databases was caused by their content in areas of medicine, physiology, or social sciences, as well as a wide selection of peer-reviewed journals (Burnham, 2006; Shultz, 2007). Search terms were a combination of the name of one of the driver state monitoring method and the terms 'circadian effect OR diurnal effect'. For example, Electroencephalography AND (circadian effect OR diurnal effect). The driver state monitoring methods used for this search were the ones previously identified in chapter 3. The search was intended for the circadian effect on the listed measures in any context, not necessarily in driver state monitoring. Abstracts of the first three hundred results were evaluated based on the inclusion and exclusion criteria and incorporated into further analysis. Table 6.1 presented the number of initial results for each search term and the number of papers that were included in the further analysis after evaluation of the abstracts.

Table 6.1: Number of results in the three databases and the number of papers selected for the analyses for every search term

	Number of initial results in Pubmed	**Number of initial results in Scopus**	**Number of initial results in Web of Science**	**Number of papers finally included in the analysis**
Electroencephalography	387	10198	45	6
Eye-tracking	2	295	4	0
Electrocardiography	303	4283	35	5
Electrooculography	29	353	4	2
Functional Near-Infrared Spectroscopy	0	562	0	0
Electrodermal Activity	6	515	7	1
Speech	15	2383	15	2
Event-Related Potential	87	12827	87	3
Electromyography	126	1938	28	1
Questionnaires	52	9147	265	1
Blood Pressure	1765	34011	2138	5
Infrared Camera	2	312	24	0
Facial Expression	10	2037	12	1
Saliva Analysis	442	5293	450	5
Body temperature	1028	26828	1806	7
Pupillometry	2	0	11	0
Respiration	300	10306	580	1
Psychomotor Performance	274	5079	208	1
Body Position	235	6382	111	1
Head Movements	5	3374	13	0
Oximetry	25	574	10	0
Actigraphy	189	4806	354	0
Blood Glucose	675	32080	529	0

	Number of initial results in Pubmed	Number of initial results in Scopus	Number of initial results in Web of Science	Number of papers finally included in the analysis
Doppler Flow Meter	0	48	6	0

Inclusion criteria. Only studies in English, with the full text available, both experimental papers and reviews were included. The papers had to describe an exact effect of the circadian phase on one of the psychophysiological functions listed before. Studies had to analyse the circadian influence on healthy, human physiological processes.

Exclusion criteria. Studies conducted on animals were excluded from the analysis. Also, studies that investigated disruptions in the circadian cycle or circadian cycles in atypical or clinical states were excluded. Papers that mentioned the effect of the circadian phase on physiology but did not describe the nature of this effect were also excluded.

As a result, 36 papers were identified and included in further analysis. Some articles investigated more than one physiological function, hence the number of the analysed papers in table 6.1 summed up to 42.

6.3 RESULTS

The evaluation of thirty-six identified papers showed that time of the day (or night) influenced psychomotor performance, body temperature, salivary cortisol, salivary alpha-amylase, speech, subjective alertness, cognitive performance, blood pressure, facial expression, ECG, EEG oscillations, ERP, EMG, EOG, EDA, respiration and bodily posture. The full list of these effects is presented in Table 6.2. A detailed description was included for the most frequently used methods of driver state monitoring. The methods were electroencephalography, eye-tracking, electrocardiography, and electrooculography.

Table 6.2: List of all the identified circadian effects on the psychophysiological measures, with references and number of papers.

Measure	Number of papers that identified the circadian effect	The type of circadian effect	References
Electroencephalography	6	Multiple changes in slow-wave sleep EEG caused by circadian phase	(Lazar et al., 2015)
		Increased wake delta power during the circadian day and decreased during the circadian night, wake alpha peak frequency decreased during the circadian night and increased during the circadian day	(Gundel & Witthöft, 1983)

Measure	Number of papers that identified the circadian effect	The type of circadian effect	References
		Increased sleep delta during the circadian night and decreased during the circadian day	(Tan et al., 2003)
		Increased alpha activity in REM during the circadian day and decreased during the circadian night, increased sleep spindles power during the circadian night and decreased during the circadian day	(Dijk, 1999)
		Nadir of alpha power around 4 a.m.	(Cajochen et al., 2002)
		Theta power highest between 4 a.m. and 8 a.m. linear decreases in the afternoon and reaches the flat plateau in the evening, lower-alpha presents a similar pattern, while the lower beta is highest between 8 a.m. and 12, and has the second peak around midnight	(Cummings et al., 2000)
Electrocardiography	5	Heart rate decreased during the circadian night and increased during the circadian day	(Gubin et al., 2017) (Prattichizzo & Galetta, 1995)
		Heart rate variability decreased in the afternoon and increased in the evening	(Cavallari et al., 2010)
		Increase of the heart rate variability in high frequencies overnight and decrease over a day	(Amirian et al., 2014)
Electrooculography	2	EOG peak-to-peak amplitude in red dot tracking task highest in the late morning with acrophase at 12:22 a.m.	(Tuunainen et al., 2001)
		An increased amount of slow eye movements during the circadian night	(Christian Cajochen et al., 1999)
Electrodermal Activity	1	A linear increase of EDA over a day	(Hot et al., 1999)
Acoustic Speech Analysis	2	A decrease of the fundamental frequency in the night	(Whitmore & Fisher, 1996a)
		An increase of the fundamental frequency in the afternoon, increase of alpha ratio and vocal loading in the afternoon for women and decrease for man	(Artkoski et al., 2002)
Event-related Potential	3	The increased amplitude of P200 in the evening in visual ERP	(Wesensten & Badia, 1992)
		Increased duration of VEP P100 and N140 components between 2 a.m. and 5 a.m. and decreased at 5 p.m.	(Stolz et al., 1988)
		P300 amplitude and latency highest in the morning	(Higuchi et al., 2000)
Electromyography	1	Elbow flexor torque has an acrophase at 6 p.m. and bathyphase at 6 a.m.	(Gauthier et al., 1996)
Questionnaires	1	Decreased subjective alertness during the circadian night, and a smaller decrease during the mid-afternoon	(Dijk et al., 1992)
Blood Pressure	5	Increased during the circadian day and 10%-20% decreased during the circadian night	(Douma & Gumz, 2017), (Gubin et al., 2017), (Hermida et al., 2004), (Prattichizzo &

Measure	Number of papers that identified the circadian effect	The type of circadian effect	References
			Galetta, 1995), (Pickering et al., 1996)
Facial Expression	1	A decrease in skin thickness, increase in elasticity and increase of wrinkles in the afternoon	(Tsukahara et al., 2004)
Saliva Analysis	5	Salivary cortisol increased during the circadian morning, decreased during the circadian night	(Del Corral et al., 2016), (Heaney et al., 2012), (Pickering et al., 1996)
		A rapid drop of salivary alpha-amylase in the morning and a gradual increase over a day	(Nater et al., 2007), (Strahler et al., 2010)
Body temperature	7	Lowest during the circadian night and highest during the circadian day, with a peak in the late afternoon and dip in the early morning	(Blatter & Cajochen, 2007), (Brown et al., 2000), (Cuesta et al., 2017), (Dijk et al., 1992), (Gubin et al., 2017), (Ekhart et al., 2018), (Gundel & Witthöft, 1983)
Respiration	1	A decrease of respiratory rate and respiratory amplitude in the night	(Bonnet et al., 1998)
Psychomotor Performance	4	A decrease in Psychomotor Vigilance Task, executive function, pre-frontal cortex related functions, sustained attention, go/no-go task during the circadian night	(Blatter & Cajochen, 2007), (Graw et al., 2004), (Sagaspe et al., 2012)
		Cognitive performance decreased during the circadian night and a smaller decrease during the mid-afternoon dip	(Dijk et al., 1992)
Body Position	1	Change in centre of pressure characteristics over the day	(Baccouch et al., 2015)

Six papers reported some circadian effect on the EEG oscillations. Both sleep and wake brain electrical activity differed depending on the circadian phase (Lazar et al., 2015). Delta power during sleep was larger in the circadian night (Tan et al., 2003), alpha activity in rapid eye-movements (REM) sleep was higher during the circadian day, and alpha spindles power was higher during the circadian night (Dijk, 1999). These differences might influence the driver's state monitoring accuracy when the driver experiences sleep episodes. In the wake EEG, delta power was higher during the circadian day (Gundel & Witthöft, 1983), theta was at its peak between 4 a.m. and 8 a.m., while beta between 8 a.m. and 12 a.m. (Cummings et al., 2000). Gundel and Witthoft (1983) and Cummings et al.

(2000) reported alpha to be at its peak in the night, while Cajohen et al. (2000) to have its nadir at 4 a.m.. Figure 6.1 showed a conceptual presentation of circadian changes in wake EEG, whilst Figure 6.2, a conceptual representation of circadian oscillations in sleep EEG. Both figures presented the visual simplification of the circadian fluctuations.

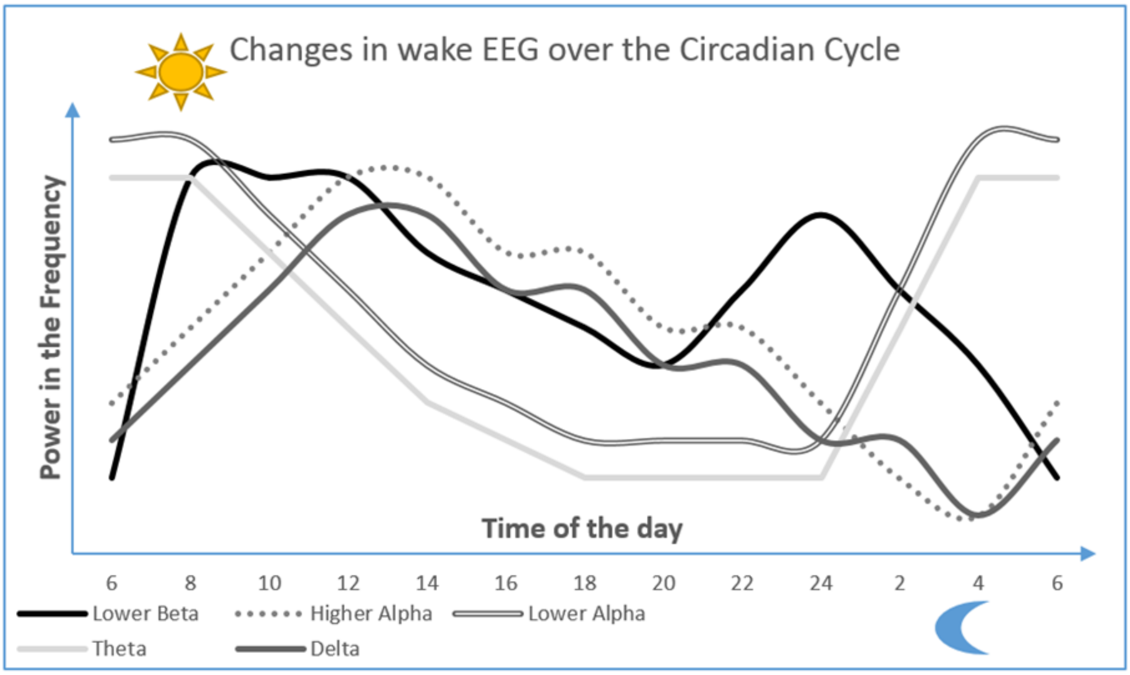

Figure 6.1: Approximate sample curves representing changes over the circadian cycle in wake EEG. Different curves represent lower beta, higher alpha, lower-alpha, theta and delta frequencies.

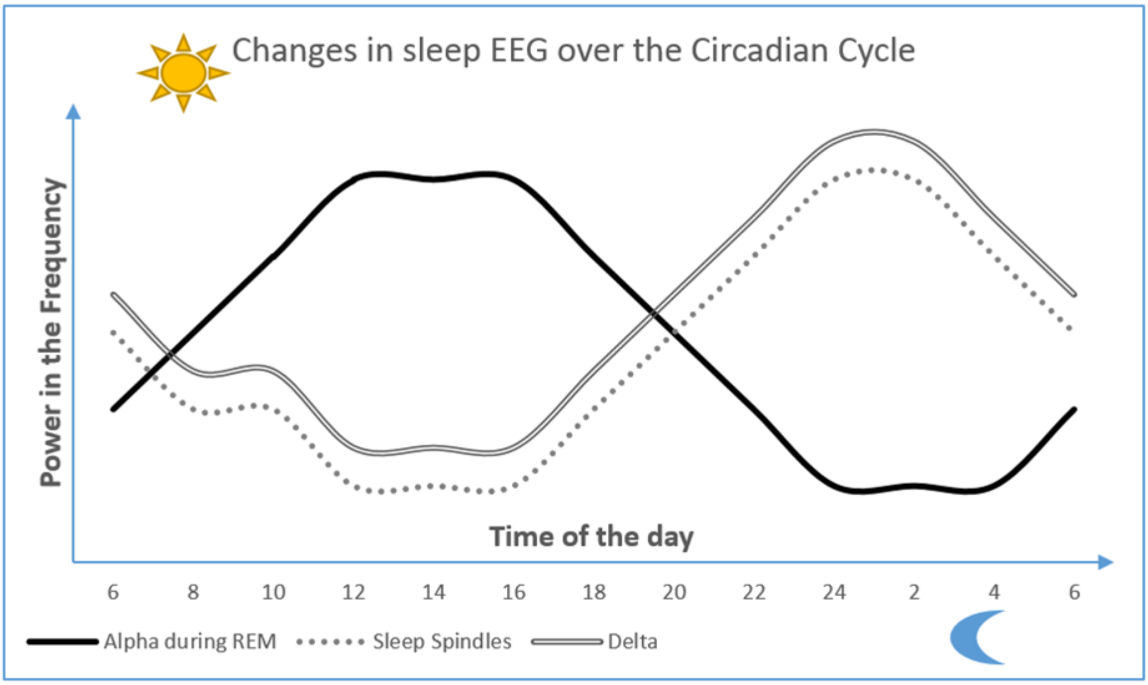

Figure 6.2: Approximate sample curves representing changes over the circadian cycle in sleep EEG. Different curves represent alpha frequency during REM sleep, power in sleep spindles and power in delta frequencies.

There were no papers on circadian effect on eye-tracking; however, two papers identified circadian changes in ocular behaviours based on electrooculography. As eye-tracking measures the same ocular behaviours (extended with the pupil size change) as an electrooculogram it could be assumed that the same circadian effects would apply to both methods. Tuunainen et al. (2001) reported that EOG peak-to-peak amplitude in the red dot tracking task was the highest in the late morning with an acrophase at 12:22 a.m. Also, an increase in slow eye movements was observed at night (Cajochen et al., 1999).

Five papers described the effect of the circadian phase on ECG. Cavallari et al. (2010) reported a decrease in heart rate variability in the afternoon followed by an increase in the evening, while Amirian et al. (2014) and Rodriguez-Colon et al. (2014) increase of heart rate variability in high frequencies over a night and a decrease over a day. Gubin et al. (2017) and Pratichizzo et al. (1995) reported that the heart rate decreased in the circadian night and increased in the circadian day. A conceptual presentation of the circadian changes in ECG was presented in Figure 6.3. It was an example, an approximate curve demonstrating heart rate and heart rate variability over the circadian cycle.

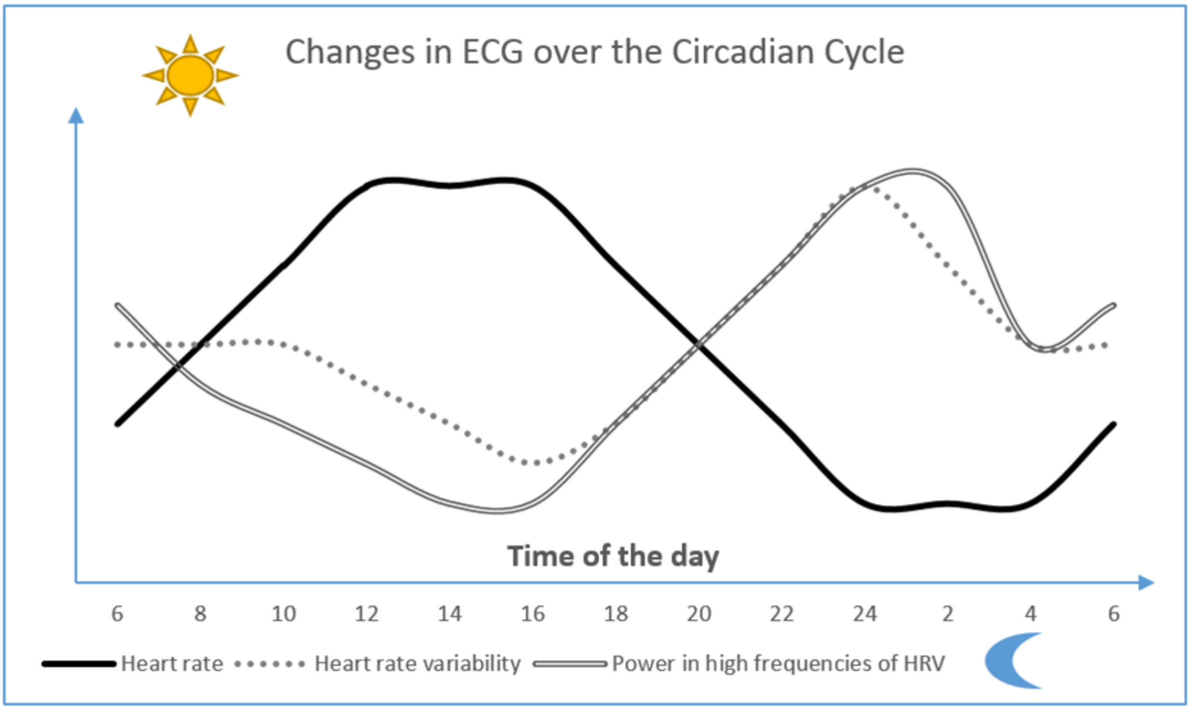

Figure 6.3: An example of the curve approximating circadian fluctuations of ECG functions. Different curves represent heart rate, heart rate variability, and power in high frequencies in heart rate variability.

6.4 DISCUSSION

This chapter reviewed the literature and identified the circadian effect on psychophysiological methods that can be used in driver state monitoring. It is important to note that increased sleepiness and fatigue could not fully explain such an effect at night. It hence would represent a different physiological baseline related to the point in the diurnal cycle.

The review was related to the measures identified in chapter 3 as potential tools of driver state monitoring. Papers that were reviewed concerned these methods but were not associated with the driving domain. They described the circadian effect on the underlying physiological processes or results obtained with these methods.

The review of the thirty-six papers showed some circadian influence on electroencephalography, electrocardiography, electrooculography, electrodermal activity, speech, event-related potential, electromyography, subjective alertness, blood pressure, facial expression, salivary hormonal content, body temperature, respiration, psychomotor performance and body position. The detailed list of all the identified circadian influences was presented in Table 6.2.

As the papers came from a variety of domains and did not specifically focus on driver state monitoring, it was not possible to assess an effect size that described circadian changes in physiology would have on the driver state monitoring. This work merely raised an issue and pointed out scientific evidence that such an effect could cause errors in the monitoring systems. The list of physiological changes due to circadian rhythmicity was not exhaustive or complete. Likely, other physiological states used for monitoring could also change depending on the time of the day (or night). For example, as body temperature changes depending on the circadian phase (Blatter & Cajochen, 2007), such change may also be visible with an infrared video camera.

The consequence of such findings is not trivial. For example, Gundel and Witthof (1983) reported that when the subject was awake power in delta frequency was higher during the day than during the night. An increase in delta power was one of the indicators of sleepiness (Dhupati et al., 2010). Let us imagine that the EEG based driver state monitoring system was used without control for the circadian phase. A certain increase in delta power would then be treated as an alarming indicator of sleepiness. However, as delta power is lower at night in wake EEG independently from the sleepiness level, the level of growth could still stay within the normal range. In such a case, the driver's sleepiness could remain undetected. It could be especially dangerous considering that night-time is a generally more difficult time for driving (Åkerstedt & Kecklund, 2001; Matthews, Ferguson, Zhou, Kosmadopoulos, et al., 2012; Mitler et al., 1988). This review identified several circadian effects that could cause changes in the baseline values and dynamics in measurements of the driver state. It was just one example of how neglecting the impact of the circadian phase could lead to a poorer driver state monitoring system. It is undeniable that the amount of literature on circadian rhythmicity in monitoring methods is scarce. Such lack of experimental knowledge could also lead to the situation when scientists and designers would not be aware of the potential consequences of the circadian phase on a particular method of driver monitoring.

This chapter created a summary of what is currently known in terms of the circadian effect on physiological mechanisms that underlie driver state monitoring methods. It used a wide variety of psychophysiological measures that were either already used in driver state monitoring, were proposed as driver state monitoring measures after laboratory tests or were used for detection of states that could jeopardize driving safety. The results led to several important conclusions. First, there is a circadian effect on many physiological functions that are used in the methods of driver state monitoring. Second, the direction and type of effect is not trivial and should be separately studied for every physiological function. The types of circadian fluctuations were described in the results section. Third, there might be a circadian effect on other physiological function that could be used in methods of driver state monitoring that were not identified in this review. It is because the reviewed research was not directly aimed to study driver state monitoring in the circadian context. It led to two conclusions, the circadian phase should be considered in the creation and design of driver state monitoring systems, and there is a need for more research in the area of circadian effect on psychophysiology, psychophysiological monitoring, and implementation of those in driver monitoring systems. This approach could also be used to monitor operators of different machines such as pilots of the aircraft, motorcyclists and nuclear plants operators to mention just a few. This research has created a theoretical basis for further investigation of the topic. It also placed chronophysiology in the area of interest of human factors engineers, designers, and safety scientists.

One limitation of this study was the small amount of literature available on the topic; however, this was also a reason why this work is a call for more research. Also, each of the analysed papers considered the physiological functions from a particular perspective, that might not necessarily apply to a driver state monitoring. For example, Tuunainen et al. (2001) observed circadian variation in the red dot tracking task, while ocular state monitoring is mostly based on blinks and saccades. Circadian fluctuation in ocular behaviours might be broader than just a red dot tracking task and further research about circadian variations in EOG and eye-tracking.

6.5 CONCLUSIONS

This review has presented a wide variety of circadian influences on the physiological functions used in the methods of driver state monitoring, namely electroencephalography, electrocardiography, electrooculography, electrodermal response, speech, event-related potential, electromyography, questionnaires, blood pressure, facial expression, hormonal salivary content, body temperature, respiration, psychomotor performance, and body position. The circadian perspective was often neglected in research in this area. It is necessary to expand existing knowledge about the circadian effect on different aspects of physiology, psychophysiological monitoring, and driving. The circadian phase should be considered as one of the essential variables in the design of the systems of driver state monitoring and driving safety. From the perspective of researchers and manufacturers, the methods of driver state monitoring should be tested and validated both during the day-time and night as some physiological functions might be less effective as indicators in different circadian phases. Also, some other safety systems should be applied for night driving. It is especially relevant in semi-automated driving as automation might increase sleepiness, and the circadian phase might influence the interaction between the automated system and the human driver (Kaduk et al., 2020). Such a perspective could also be relevant for operators of other machines, their monitoring and the general topic of the shift work.

7. CREATION OF THE LABORATORY FOR PSYCHOPHYSIOLOGICAL MEASUREMENTS OF THE DRIVER

Psychophysiology offers great potential to human factors engineering and ergonomics (Parasuraman, 2011); however, at the same time, it introduces some challenges related to the proper measurements of the physiological functions. This chapter described the process of developing a psychophysiological laboratory with a driving simulator. The challenges of the laboratory construction that aimed to reduce the amount of noise in signals, other methods of noise reduction, choice of optimal measures, and selection of brands of the psychophysiological devices were also described.

7.1 CHOICE OF THE MEASURES

Literature review in chapter 3 allowed to identify several methods as potentially useful for the driver's state identification. EEG with analysis of oscillations, eye-tracking, ECG, EOG, functional near-infrared spectroscopy (fNIRS), EDA, acoustic speech analysis, ERP (event-related potential), EMG, questionnaires, blood pressure, infrared video camera, facial expression, saliva analysis, body temperature, pupillometry, respiration, driving performance, psychomotor performance, head movements, oximetry, actigraphy, blood glucose, and Doppler flow meter. Due to the technical, temporal and financial limitations, there was a need to select the most promising measures for further experimental evaluation.

Initially, some methods were eliminated due to their low usability. Blood pressure measure was eliminated due to its low specificity (Roscoe, 1992), Doppler flow meter due to lack of experimental literature that would prove its effectivity in the state identification, and the psychomotor tests because performing them would highly interfere with the driving tasks. Body position, head movement and facial expression measures were also excluded, because they present a high challenge for computer vision and tactile recognition systems, while the study was mostly concentrated on psychophysiology (Fan et al., 2010; Murata et al., 2015). Blood glucose was excluded because of the low number of citations and high invasiveness of this measure. Eye-tracker was initially included in the experiment because the necessary equipment was already in possession of the laboratory; however, the technical problems led to further exclusion of this tool. As eye-tracker offers the possibility of pupils' size measure, pupilometer was treated as not necessary and excluded.

The tools necessary for ECG, respiration and blood oxygenation measurement were already in possession of the human factors group, so they were arbitrarily included in the study.

The questionnaires were included because they did not bring any additional financial costs and could provide important information about the level of self-awareness that people have about their state while driving.

To choose the methods from the rest of the measures, a systematic cost versus benefits analysis was conducted. The results of the analysis were presented in Table 7.1. A point system was created to evaluate the costs versus benefits balance for each method. The following criteria were used: spatial accuracy, temporal accuracy, the accuracy of the state classification reported in the literature, prevalence of the papers using this method in the driving safety research, prevalence of the papers using this method in the state detection research outside of the driving domain, financial cost of the necessary devices, and susceptibility to artefacts. The arbitrary system of points was used. Positive points were in the range of 0 to 3, and negative points in the range of 0 to -3. They expressed the usability of the method. A higher score expressed higher usability.

Table 7.1: Summary of the points that psychophysiological measures received for different usability criteria and the final usability classification.

	EDA	**EEG**	**fNIRS**	**EOG**	**Acoustic Speech Analysis**	**Actigraphy**	**EMG**	**Saliva Analysis**	**Near-Infrared Camera**
Spatial Accuracy	2	1	2	3	2	2	3	2	3
Temporal Accuracy	3	3	2	3	2	2	3	0	3
Classification Accuracy	2	3	2	2	3	3	2	2	2
Prevalence in Driving Papers	2	3	1	3	1	1	1	1	1
Prevalence in State Classification Papers	2	3	2	2	1	0	1	1	1
Novelty	0	0	1	0	0	0	0	0	0
Cost	0	-1	-3	0	0	0	0	-1	-2
Susceptibility to artefacts	-3	-3	-3	-2	-1	-1	-2	0	-1
Sum	8	9	4	11	8	7	8	5	7
Usability Score	High	High	Medium	High	High	Medium	High	Medium	Medium

Spatial and temporal accuracy were assessed arbitrary based on Parasurman (2011), and Parasurman et al. (2008), as well as the expertise of the researcher (with 0 points for the lowest accuracy, and 3 for the highest accuracy). As temporal and spatial accuracy are related to many factors, there was a need for an arbitrary numerical value to express the accuracy of the method. Temporal accuracy was treated as very high when it was possible to obtain a physiological gauge after 1-5 second and as very low when it was necessary to wait for the result for more than two minutes. In the case when the accuracy did not apply to the method, the score was two.

The accuracy of the state classification was averaged between papers from the review presented in Chapter 3. It was only considered when accuracy was unambiguously mentioned in the paper: with 0 for the accuracy lower than 40%, 1: 41%-60%, 2: 61%-80%, 3: 81%-100%. If the method was used in combination with other methods with accuracy higher than 90% more than three times, one accuracy point was added to the score.

Prevalence of the method in the driving safety papers was based on the literature review (see Chapter 3) and expressed in the following way: 0: 0 papers, 1: 1-5 papers, 2: 6-10 papers, 3: more than 10 papers. Prevalence of the method in the state classification papers was based on the literature review (see Chapter 3) and expressed in the following way: 0- 0 papers, 1: 1-5 papers, 2: 6-10 papers, 3: more than 10 papers. Papers with high (2 or 3 points) prevalence of method in state classification, but low (0 or 1 points) in driving safety classification received one additional point for the potential for the novelty of the findings.

Negative points were allocated for the costs based on the cheapest quote given by the companies in the following way: 0: less than 5000 pounds, -1: 5001-15 000 pounds, -2: 15 001- 25 000 pounds, -3: more than 25 000 pounds.

Susceptibility to the artefacts received -1 point for susceptibility to the muscular movement like blinking, breathing or talking, -1 for susceptibility to the general body movement like walking, typing etc. and -1 point for susceptibility to the external artefacts like noise, light or temperature changes (Cacioppo et al., 2007; Mehta & Parasuraman, 2013; Parasuraman, 2011; Wickens et al., 2015).

The final score was classified as an indicator of usability, with the following interpretation: less than 4 points- small usability, 4-7 points: medium usability, 8-11points: high usability, and more than 11 points: very high usability. According to the proposed classification, EDA, EEG, EOG, acoustic speech analysis and EMG were classified as high usability methods. Actigraphy, saliva analysis, infrared camera and fNIRS as the medium usability methods.

This method of the usability assessment was created as a tool for this research and presented just a part of information related to the value of the assessed research methods. It was used as a strategy of measures selection for this study.

Out of the measures classified as highly usable, all were included in the experiment. Additionally, saliva cortisol and alpha-amylase analysis were included in the category of medium usability methods. The decision was based on the costs (fNIRS and infrared camera costs were higher than the available funds). The second reason was the fact that the analysis of saliva introduced an entirely new category of measures to the experiment. While most of the measures controlled electrical activity of different body parts, saliva analysis is a hormonal measure that might give a new perspective of the physiology of driving. Additionally, both cortisol and alpha-amylase are strongly influenced by the circadian rhythms and cortisol is sometimes used as a circadian phase indicator, which gives another tool for the circadian analysis in the study (Del Corral et al., 2016; Strahler, Berndt, Kirschbaum, & Rohleder, 2010).

Even though EEG was selected as a tool, the results of EEG were not presented in this thesis due to the temporal restraints, technical reasons, and highly time-demanding pre-processing process. Nevertheless, the data related to the noise reduction in EEG recording as well as EEG brand choice was included in this chapter as it might present a value of the technical recommendation.

7.2 CHOICE OF THE BRANDS

The market offers a variety of models and brands of psychophysiological measurements devices. A review of the available models was conducted to assure the most beneficial choice.

As the research group already possessed ECG, and respiration measurement devices from BioPac, purchasing EDA, EMG and EOG from this company allowed the lowest price and possibility to combine modules and use the software that was already available.

A microphone, pre-amplifier and a recorder for speech analysis were chosen based on several indicators. The characteristics of the devices were based on the scientific literature. The microphone was selected based on the frequency curve (it should be flat). Recorder and pre-amplifier were chosen in the way to preserve recorded sound from frequency modifications (Švec et al., 2003; Hunter et al., 1997).

The choice of the laboratory engaged in saliva analysis was based on the Salimetrics approved laboratories list. The only laboratory in the United Kingdom that had such approval was Anglia Ruskin BioLab and it was contracted for the saliva analysis.

The biggest challenge was presented by the choice of the EEG device model. Four offers met the financial constraints of the project Mobita 32, Enobio 20, EEGO Sports and Quick 20. They were all wireless, portable models. Mobita 32 had water tap electrodes, Quick 20 dry electrodes, EEGO Sports gel electrodes, and Enobio 20 offered both dry and gel electrodes. As dry electrodes were reported to generate more signal noise (Mathewson et al., 2017), Quick 20 was excluded. Experimental papers using three remaining models: Mobita 32, Enobio 20 and EEGO Sport were analysed from the point of view of signal noise, mobility, quality of the signal and use in driving research. The effect of analysis for each of the models, together with the list of investigated research papers can be found in the tables in Appendix 4. The summary of the comparisons was presented in Table 7.2.

Table 7.2: Summary of the data related to the 3 analysed EEG models Mobita 32, Enobio 20 and EEGO Sports.

EEG Model	**Amount of papers**	**Data quality**	**Level of Mobility**	**Number of channels**	**Other comments**
Mobita	4	Very good, but only one paper analysed it	High	32	Shielded electrode cables
Enobio	>12	Very good in comparison with other EEG models, used multiple times in BCI projects with high accuracy	High	20	
EEGO Sport	>7	No actual analysis of data quality, but one study managed to measure EEG oscillations during running and walking, another obtained good quality P300 outside of the lab, used multiple times with high accuracy of classification	High	32	

As a result of this analysis, Enobio 20 was chosen for the purchase due to the high number of experiments that used this model with good data quality and a high number of driving research papers. All the papers were listed in the table in Appendix 4.

7.3 CONSTRUCTION OF THE LABORATORY

7.3.1 MONITORING WITH MULTIPLE SENSORS

The development of automated technologies made monitoring the operator's state more important to ensure safety. One of the examples is semi-automated driving, where the function of the driver is shifted into a more supervisory role (Kyriakidis et al., 2019). Unfortunately, people tend to perform poorly in tasks that require sustained attention (Warm et al., 2008).

Driver state monitoring could improve safety by surveillance of a driver to ensure that they are in an appropriate state and engaged in driving-related tasks (Kyriakidis et al., 2019). The combination of multiple physiological recordings might provide more information. Moreover, a comparison of effectivity and accuracy of different measures might enable the most optimal choice of the monitoring system. It could allow a selection of the method most suitable for requirements of a particular situation in a manner of cost, efficacy, accuracy, speed and others. However, conducting a study using multiple sensors creates specific challenges. Each type of physiological signal might be confounded by the different kinds of noise that should be reduced as far as possible (Sweeney et al., 2012). A combination of various measures might increase the number of potential sources of noise that can affect such measurements (Cacioppo et al., 2007; Sweeney et al., 2012). The control of all potential signal issues across methods requires a variety of strategies, which can increase the complexity of the experimental environment. The processes undertaken are often needed to become more complex and cumbersome. A combination of different measures might also be affected by a combination of various noise sources; therefore, might require more complex noise control. Additionally, a set-up of multiple sensors can be more complicated, especially if they interfere with each other.

There is a gap in the literature regarding the set-up and the laboratory space preparation that would ensure the best data quality for the multiple simultaneous psychophysiological recordings. The aim of this chapter was to address this gap by describing the experimental set-up and laboratory construction for a multisensory recording on the example of a multisensory, low-low fidelity driving simulator. It provided a decision-tree with some recommendations. The unique suite of recommendations was generated based on a specific set of methods chosen by a user. It is meant to be a prototype of a practical tool for enhancing high-quality data collection.

7.3.2 MULTIPHYSIOLOGICAL MEASUREMENTS IN HI:DAVE

This research project studied driver state monitoring using a wide variety of different psycho-physiological measures. Initially, proposed methods were electroencephalography (EEG), camera-based eye-tracking, electrooculography (EOG), electromyography (EMG), electrocardiography (ECG), respiration belt, electrodermal response (EDA), saliva-based cortisol and alpha-amylase analysis, acoustic voice analysis, oximetry and questionnaires. These measures were chosen after the completion of an extensive literature review revealing what measures have the maturity for use in applied environments in the short, medium and long term (see Chapters 3, 7.1, and 7.2). After the pilot study, the eye-tracker was removed from the measurement methods due to its interference with forehead electrodes of EEG, EMG, and EOG and high computational load of the recording, such a variety of the recorded signals required a cautious approach towards noise reduction. So-called signal noise, which is a recording of artefacts, can be reduced with an appropriate laboratory construction and experimental set-up. Unavoidable noise can be rejected from the data with different strategies of data pre-processing. However, algorithms that reject artefacts may lead to some signal loss; therefore, the experimental set-up should be designed in a manner to reduce the artefacts as far as possible. Two versions of the experimental set-up were depicted in figure 7.1 and 7.2.

Figure 7.1: Participant driving a driving simulator with EOG electrodes around the eyes, EMG electrodes on the forehead, EDA electrodes and oximetry sensor on the fingers. It is also visible that he has a respiration belt around the abdomen and a head-mounted microphone. The ECG electrodes are hidden under the t-shirt.

Figure 7.2: : Participant driving a driving simulator with EEG cap on the head, EDA electrodes and oximetry sensor on the fingers. It is also visible that he has a respiration belt around the abdomen and a head-mounted microphone. The ECG electrodes are hidden under the t-shirt.

7.3.3 TYPES OF SENSORS AND SOURCES OF NOISE

EEG, EMG, EOG, and ECG estimate an electric activity of the brain, muscles, retina and heart, respectively. They use electrodes placed on the surface of the skin (Chang, 2010; Phinyomark et al., 2012; Reddy et al., 2011; Urigüen & Garcia-Zapirain, 2015). The EDA is a measure of electric potential on the skin surface assessed between two electrodes (Taylor et al., 2015). Eye-tracking is a broad group of methods of eye movements monitoring (Duchowski & Duchowski, 2017). In the case of this research project, considered eye-tracker was a helmet with two cameras recording the eyes and one camera recording the visual field; however, there is also an option of eyes monitoring with desktop mounted camera (Duchowski & Duchowski, 2017). There is a variety of devices measuring breathing amplitude and frequency, but in this project, it was measured with a belt placed on the upper abdomen (Sweeney et al., 2012). Acoustic voice analysis uses a voice recording to analyse the acoustic characteristics of the speech (Plichta, 2002). Oximetry is the measure of blood oxygenation level and a pulse with the use of near-infrared light generated by a device placed on a finger or an ear (Ram et al., 2012; Sweeney et al., 2012).

Each of the methods above has a specific set of potential artefact sources that need to be controlled to achieve as clean data as possible. In the case of acoustic speech analysis artefacts in the signal might be either caused by the noises other than speech or by changes in signal property caused by the microphone, pre-amplifier, recorder or data conversion (Plichta, 2002). Electrophysiology is sensitive to any type of movements, muscle activity, electrodes displacement, sweat, and influence of surrounding electromagnetic signal, even the signal of the recording devices (Kirst et al., 2011; Phinyomark et al., 2012; Rahman et al., 2011; Sweeney et al., 2012). Likewise, the measurement of the electrodermal activity can be disturbed by electromagnetic impulses, body movements, temperature changes or the displacement of the electrode (Taylor et al., 2015). Respiratory signals can be changed with body movements (Sweeney et al., 2012). The eye-tracking signal can be disturbed by blinking (in the case of gaze monitoring), eye-lashes and light instability (Duchowski & Duchowski, 2017). Oximetry is susceptible to movement, other body signals, electromagnetic influences (Chong et al., 2014) and changes in the light (Sweeney et al., 2012). Many of the noise sources are quite common in the environment; for example, electromagnetic impulses are generated by the surrounding electrical devices or electrical wires. Similarly, acoustic noise can be generated by objects like fans, crackling noise from the cables, elevators and others.

7.3.4 EXPERIMENTAL SET-UP: A DECISION TREE

The decision tree presented below was a result of a literature evaluation and hands-on comparison of different set-up strategies. It is a prototype of a guide tool that could support the choice of the most optimal experimental set-up for the measures selected for the experiment. This tool was an attempt to answer the gap in the literature that would combine recommendations for set-up strategies while combining different physiological measuring devices. It included EEG, EMG, EOG, ECG, EDA, camera-based eye-tracking, respiration belt, oximetry and voice recording. The user should follow descending decision nodes and answer the questions asked in the nodes. The arcs represent the answers ‘yes’ and ‘no’. They lead to additional questions and finally to the unique set of recommendations for the particular experiment. Recommendations were expressed as a list of numbers. Numbers were explained in table 7.3. The explication of the recommendations was included in the next chapters. Due to the complexity of the decision-tree and the visibility issues, the graph was divided into two parts 2A and 2B. A user following the arcs from graph A to graph B should continue following the arc with the same numerical description.

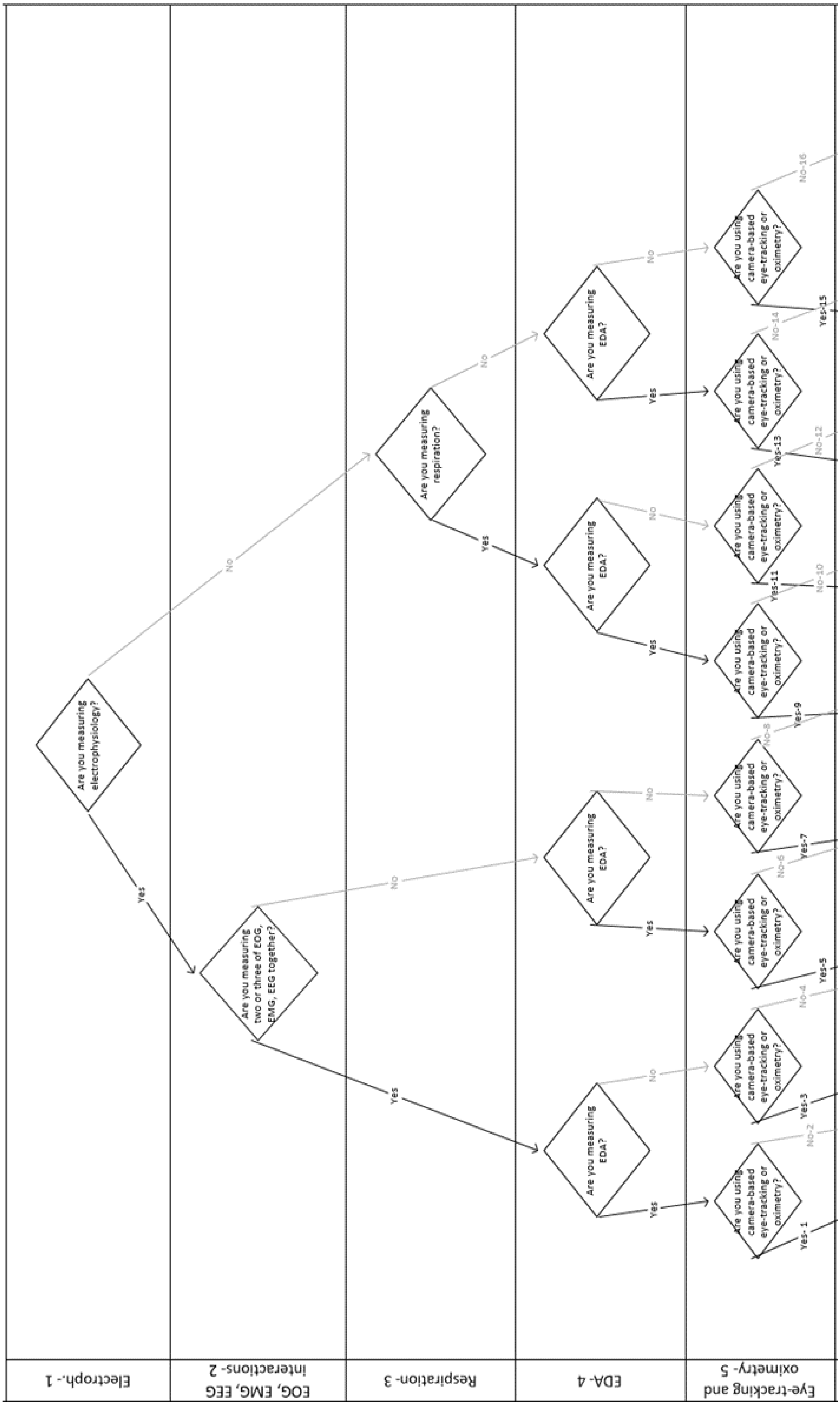
Electroph.- 1
EOG, EMG, EEG interactions- 2
Respiration- 3
EDA- 4
Eye-tracking and oximetry- 5
Are you measuring electrophysiology?
Are you measuring two or three of EOG, EMG, EEG together?
Are you measuring respiration?
Are you measuring EDA?
Are you using camera-based eye-tracking or oximetry?
Yes
No

Fig. 2 A

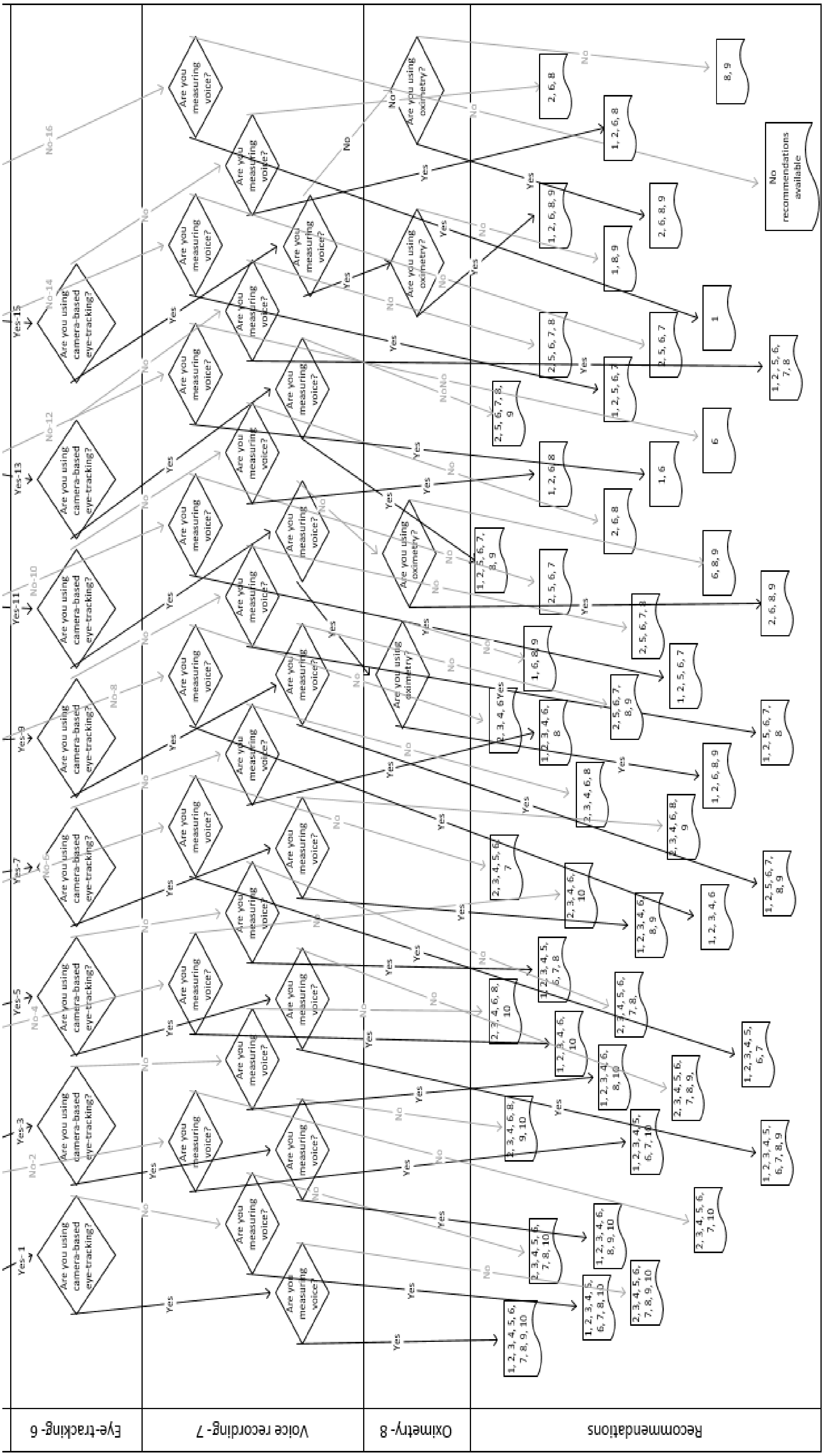
Eye-tracking- 6
Voice recording- 7
Oximetry- 8
Recommendations

Fig. 2 B

Figure 7.3: The flow chart presents a decision-tree leading to some recommendations for noise reduction in an experiment. The chart contains two parts, figure 2a and figure 2b. Each level represented one decision, for example, if electrophysiology is used in the experiment. A black connector on the left represented a ‘yes' choice and a grey connector on the right represented a ‘no' choice. The bottom level consisted of a unique set of recommendations for the chosen experimental design. Numbers representing recommendations werelisted in Table 7.3

Table 7.3: Recommendations and representing numbers from the decision tree (Figure 7.3).

Number	Recommendation
1	Reduce acoustic noise in the lab, insulate walls from noise, use silent equipment.
2	Reduce electromagnetic influences, preferably use a Faraday Cage.
3	Prepare the skin for the electrophysiological electrodes.
4	Use the electrolyte for the electrophysiological electrodes.
5	Ask participants not to use soap, detergents, alcohol or hand-cream in the area where you apply EDA electrodes.
6	Reduce the movements of the participant. 7
7	Keep a stable temperature in the lab.
8	Keep the light in the lab at a stable level.
9	Ask participants not to use mascara.
10	Avoid the interactions between facial EMG, EOG and EEG electrodes, choose an alternative montage of EOG.

7.3.5 ARTEFACTS REDUCTION- EXPERIMENTAL SET-UP

Some techniques exist to reduce noise and artefacts contained in physiological data(e.g. artefact detection algorithms). However, the accuracy of such tools is not always apparent and the process typically leads to a loss/replacement of data, which can negatively impact the analysis results once such pre-processing has been completed. The best way to ensure high-quality physiological recording is via a diligent experimental set-up and data collection process, reducing the level of noise contained in the raw data collected (Plichta, 2002).

ELECTROENCEPHALOGRAPHY

To avoid movement artefacts, participants are often required to fixate their gaze (Plöchl et al., 2012) and reduce movements (Islam et al., 2016); however, it is hardly feasible in many experimental scenarios. For instance, driving tasks might require hand movements, whole-body movements and vibration. Due to such difficulty, EEG devices are often used in combination with additional instruments to measure potential sources of signal disturbances, like EOG, EMG, gyroscope (O’Regan et al., 2013), eye-tracker (Noureddin et al., 2012) or a camera recording head movements (Bang et al., 2013). In such instances, it is vital to avoid putting electrodes too close to each other or pressing EOG/EMG electrodes with an EEG cap. Such pressure could lead to a detachment of the electrodes and increased skin impedance. Any detachment of the electrodes or deficiency in the electrolyte can decrease signal quality (Nolan et al., 2010). Other factors might also increase skin impedance, for example, sweat or not enough electrolyte used. To decrease the impedance of the skin, there should be a thorough skin abrasion applied to remove a dead epidermis (Burbank & Webster, 1978). Impedance and signal quality might also depend on the equipment used. Even though there is a constant development of EEG technologies, devices recording signal with dry electrodes provide significantly worse data quality than gel or tap water-based. Commercially distributed, cheap EEG devices with a small number of dry electrodes, also have significantly worse quality than scientific models (Pinegger et al., 2016). In terms of the electrodes, montage caps give the higher signal quality that headsets supporting electrodes (Bang et al., 2013).

Electroencephalography was used in the experimental work for this PhD; however, due to the temporal restriction, technical problems with the recorded signal, and high time-demand of the data pre-processing, the analysis of EEG was not included in the results section.

EYE-TRACKING

Most of the video-based eye-trackers require a calibration process before the recording and repeated during the experiment, which can be time-consuming and break immersion during the experimental procedure. Some authors recommend planning a short experiment to avoid repetitive calibrations (Duchowski & Duchowski, 2017). Aside from this, some simple steps can reduce the likelihood of common signal disturbances. For example, eyelashes are often a problem; participants should not wear mascara for the experiment. A further consideration is that the laboratory should have a stable light intensity (Duchowski & Duchowski, 2017). In the case of this research project, eye-tracker was significantly reducing signal quality from the forehead electrodes of EOG and EEG and appeared to cause headaches in the pilot participants. The computational weight of the recorded signal led to frequent crashes of the software. Due to these problems, an eye-tracker was not used in the experiment.

ELECTROOCULOGRAPHY

EOG is often measured together with EEG or EMG. Many EEG devices include additional EOG electrodes that use the same grounding and reference as EEG. However, if a separate EOG device is used, electrodes for vertical eye-movements might interfere with forehead electrodes or a cap in EEG. An additional challenge is the placement of ground electrodes that is mostly recommended on the bony area in the middle of the forehead (Reddy et al., 2011). These electrodes can interfere with EEG electrodes, EEG cap and the electrodes of the forehead EMG. Therefore, it is recommended to use those devices separately. If they need to be combined, it is acceptable to use alternative EOG electrodes placement with two electrodes placed in the external eye edges and only one ground electrode in the middle of the forehead (Lopez et al., 2016). Previous work described experimentally validated, acceptable set-ups of the EOG electrodes (Lopez et al., 2016). Same as with other electrophysiological methods, an electrode site should be prepared with a thorough skin abrasion process (Burbank & Webster, 1978). In this experiment, EEG and EOG were used separately due to the interference between the electrodes. Participants were randomly assigned to the group with EEG measurement or EOG and forehead EMG measurement.

ELECTROMYOGRAPHY

EMG electrodes can be placed anywhere on the skin surface. It is recommended not to combine forehead EMG and frontal cortex EEG or an EEG cap. In this project, it was observed that signal quality decreased and impedance increased more rapidly over time when forehead EMG electrodes were additionally pressed by the EEG cap. In a case, when EEG and forehead EMG need to be combined, the solution might be using a loose electrodes EEG montage. As with other electrophysiological methods, an electrode site should be prepared with a thorough skin abrasion process (Burbank & Webster, 1978).

ELECTROCARDIOGRAPHY

The skin impedance highly influences the quality of the signal received by electrodes. Therefore, it is recommended to use wet electrodes with highly conductive electrolyte to maximally reduce the impedance (Kirst et al., 2011). Same as with other electrophysiological methods, an electrode site should be prepared with a thorough skin abrasion process (Burbank & Webster, 1978).

RESPIRATION

Respiration measured with an abdominal belt is quite an artefact-resilient method. However, like many other physiological signals, it can be disturbed by the noise created by body movements. Hence, it is recommended that the participant remains still during the experiment (Sweeney et al., 2012). In this research project, it was observed that when participants wore the top of a slippery fabric, it often led to belt displacement.

ELECTRODERMAL RESPONSE

As EDA is a measurement related to the sweat glands activity, the laboratory that uses this method should put special attention into keeping a stable temperature in the room (Measures et al., 2012). Different guides recommend placing EDA electrodes on the non-dominant hand if the measurement is taken from the fingers, palm, or wrist. Unlike electrophysiology, the skin should not be scraped before the recording. The use of soaps, alcoholic substances and other detergents can disturb the recording and because of that participants should be asked to clean their hands only with water before the experiment and not to use the hand-cream (Cacioppo et al., 2007).

ACOUSTIC SPEECH ANALYSIS

External noise can be reduced with a noise-insulation of the lab space (Plichta, 2002). It will be described in details in subchapter 7.3.6. Other research devices that need to stay inside the insulated room might produce low-frequency noise, for example, with fans or cables buzzing. If possible, the choice of the other devices should take into account the level of the low-frequency noise that they generate (Plichta, 2002). Especially the recording equipment should be as noise-free as possible. The pre-amplifier should have a balanced XLR to minimalize artefacts caused by cables. It should, also, have a high gain, broad dynamic range, high SNR (signal to noise ratio), and phantom power (Plichta, 2002). The microphone should be either head-mounted or kept in the stable, very close distance from the mouth. The distance of the four centimetres is preferable and allows to reduce lots of external noise (Plichta, 2002). It is important to choose a microphone that has a wide and flat frequency curve to avoid different responses to the different frequencies of the speech (Plichta, 2002). Omnidirectional microphones mostly have a more even response to different frequencies; however, they should be used only in very quiet laboratories (Hunter et al., 1997).

OXIMETRY

Oximetry is an optical method; hence, it is susceptible to the noise related to sources of light. It is recommended to keep the light at a constant level. Movements of the participant can also disturb the measurement, so they should be reduced to the minimum (Ram et al., 2012; Sweeney et al., 2012). If the sensor is placed on the finger, it is beneficial to put it on the less active hand. It is not recommended to use finger sensors on the ears and vice versa, due to a decrease in signal quality (Haynes, 2007).

7.3.6 ARTEFACTS REDUCTION- LAB CONSTRUCTION

The laboratory environment for a voice recording should be noise-insulated. Some researchers use anechoic chambers (Hunter et al., 1997); however, they are expensive and pre-dominantly used solely for voice recording purposes. In the case of multisensory recording, the laboratory should meet multiple requirements to reduce different noise types. Therefore, it is optimal to build an isolation booth with different types of insulation depending upon the sensory recording(s) being focused upon. If recording voice, for example, the booth, should be constructed using materials that provide noise insulation, ideally with double walls or walls insulated with acoustic material, and a floating floor (Hunter et al., 1997). It is optimal to remove loud devices from the booth, such as air-conditioning fans, loud lights and PCs with loud fans (Hunter et al., 1997).

In an experiment using electrophysiological methods, electrodermal activity, or oximetry it is beneficial to reduce the surrounding electromagnetic signals (Chong et al., 2014; Kirst et al., 2011; Taylor et al., 2015). The recommended method is the construction of the Faraday cage (Fathima & Umarani, 2016).

Additionally, in the case of EDA measurements, the temperature inside the booth should be kept at a stable level (Measures et al., 2012).

The optical measuring methods, such as oximetry and camera-based eye-trackers, are susceptible to sources of light and because of this, it is recommended to keep the light inside the laboratory at a constant level (Duchowski & Duchowski, 2017; Sweeney et al., 2012).

The laboratory constructed for this project was a noise-insulated Faraday Cage configured with a low-fidelity driving simulator, as shown in figures 7.4, 7.5, and 7.6. However, it would be possible to reconfigure the simulation suite to facilitate studies using physiological recordings across many domains (e.g. flight simulator), as the fundamental issues, concerning the quality of data collection would remain the same. The walls inside the booth were covered with multiple levels of heavy-duty aluminium foil (Fathima & Umarani, 2016) and later with fire-retardant plastic to reduce light effects caused by the aluminium foil. The booth was constructed from plywood and a fire-retardant cortex-like plastic. PCs, BioPac signal receiver and an oximeter were placed outside of the booth to reduce their electromagnetic influence on the signal. The holes were drilled in the walls to put receiving antennas and power cables through them. The receiving antennas from the

BioPac receiver were extended with RP-SMA cables and the Enobio-20 USB receiver was extended with a USB extension cable. Power cables and a voice recorder that had to stay inside were wrapped in aluminium foil.

Figure 7.4: Outside of the laboratory booth insulated with fire-retardant acoustic foam to reduce acoustic noise in the voice recording.

Figure 7.5: Inside of the laboratory booth insulated from electromagnetic noise with several layers of a heavy-duty aluminium foil.

Figure 7.6: The way to connect aluminium foil with a cable. Cable has only a ground wire inside. The cable is led outside of the booth through the hole drilled in the wall and switched to the socket.

To ensure a reduction in electromagnetic interference, the Faraday Cage had to be evenly covered with the conductive material. It does not have to be a sealed unit; it can also be covered with mesh or wire with the size of holes respective to the level of undesired frequencies. Materials recommended for electrophysiology are copper or several layers of heavy-duty aluminium foil (Cutmore & James, 1999; Fathima & Umarani, 2016). Even though small cages can properly shield without a grounding, much larger Faraday Cages need to be grounded to maintain shielding properties (Cutmore & James, 1999). In the case of this Faraday Cage aluminium foil from two walls were formed into ‘pony-tails’, clipped with a metal clip was used to connect these to the cables. The cables were plugged into the electrical sockets outside of the laboratory booth for grounding purposes. The cables were led outside through the holes drilled in the walls.

Touching the aluminium walls of the Faraday Cage might cause a tiny shock due to the static potential being transported from the body or clothes to the ground. Therefore, it is recommended to put non-conductive shielding on the walls. It is, also, beneficial as it reduces light reflexes created by the aluminium that might disturb eye-tracking recording.

It is important to remember that all of the materials (e.g. staples) used to attach things to the walls should be highly conductive, so they do not disturb the Faraday Cage effect. In the case of this project, plastic sheets were attached to the aluminium walls with construction staples made of metal.

This chapter presented a decision tree, which could be utilized when undertaking the construction of a space to be used for applied experimentation using physiological recording techniques. The construction of a noise insulated faraday caged laboratory was detailed alongside recommendation for signal noise reductions.

8. TIME-COURSE OF SEMI-AUTOMATED DRIVING- EXPERIMENTAL RESULTS

8.1 INTRODUCTION

The literature presented in the previous chapters raised some concerns related to the safety of semi-automated driving and especially a decrease in driving performance related to the automated mode. At the same time, there were some predictions related to the positive influence of automation on driving safety. Automation might offer the potential to mitigate or reduce driving risks (Stanton & Marsden, 1996). It can also increase efficiency, alleviate the workload, and improve transport capacity (Kyriakidis et al., 2019). In the aviation domain, the use of automation has increased over the past three decades and has contributed significantly to improved safety (Chialastri, 2012). Therefore, it is anticipated that higher levels of automation can be incorporated into automobiles to reduce safety risks to road users and pedestrians alike (Kyriakidis et al., 2019). There are, however, still concerns related to the role of a human driver in partially automated systems. Even though some functions might be automatized, there is an ongoing discussion about who would be responsible for the system's failure. Similarly, the ability of human drivers to safely interact with automation is questionable (Hancock, 2019). As a result, along with all the benefits, semi-automated driving introduces a specific set of challenges. For example, staying attentive while using automatic driving mode requires sustained attention and vigilance. People find it tiring, hard, and stressful to stay vigilant for longer periods, especially during monotonous tasks (Hancock, 2015; Warm et al., 2008). In a perfect world with no automation failures, the driver may still be required to monitor the driving processes but would not necessarily have to intervene. Such a situation would require long periods of attention but without stimulating tasks, which could help the driver to maintain vigilance. As such, while semi-automated driving requires vigilance, it also creates a difficult environment for its maintenance. Even if drivers restrained themselves from activities not related to driving (e. g., texting or reading), they might still experience cognitive distraction (Liang & Lee, 2010), fatigue or increased sleepiness (Schömig et al., 2015; Warm et al., 2008). However, not all the experimental data confirm the concerns related to automated mode. Some show no performance decrease in high-automation (Merat et al., 2014). Until level 5 automation is released; however, there will always be a requirement for the driver to both monitor the automation and to take back control at some point (Kyriakidis et al., 2019; Warm et al., 2008; Young & Stanton, 2002). It is related to the fact that automation of the variety of driving functions does not necessarily introduce autonomy of the vehicle.

Another potential challenge is suboptimal mental workload. The data show an inverted U-shaped association between mental workload and performance. If automation reduced the amount of mental workload, it could lead to underload (Heikoop et al., 2016; Young & Stanton, 2002, 2002, 2007), and as a result, decrease driving performance. Decreased performance could mean worse or slower intervention as well as worse driving performance after take-over. Paradoxically, a less demanding automated phase could increase the driving risk.

Another risk could be related to night-time semi-automated driving. Accident risk significantly increases at night during manual driving (Matthews, Ferguson, Zhou, Sargent, et al., 2012; Mitler et al., 1988). Based on the review of the literature, it has been proposed that such risk might be even more pronounced in semi-automated driving (Kaduk et al., 2020), possibly because of circadian changes in fatigue and attention. The increased demand for sustained attention might induce fatigue at any time of the day or night, but night-time is generally related to increased fatigue. Likewise, sustained attention is one of the functions highly affected by circadian rhythmicity and it gets worse during the night (Kaduk et al., 2020).

Owing to the monitoring requirement, the need to take-over and the possible need to intervene remains an integral part of the semi-automated system not only as a user but also as an active agent. Because of this, proper human-automation interaction remains crucial for driving safety (Kaber, 2018). It is therefore vital to understand the effect that automation has on manual driving and driver state.

This chapter presented the results of the analysis of the experimental data related to the time-course of semi-automated driving, the dynamics of the driver performance, driver state, and driver physiology. The hypothesis of this chapter was that there is a change in driver performance and driver state after automated mode characterised by a decrease in driving performance and a less alert state.

8.2 EXPERIMENTAL METHODS

8.2.1 PARTICIPANTS

CALCULATION OF THE SAMPLE SIZE

The data analysis used a variety of different statistical tools; however, the required number of participants was calculated based on the assumption that the main statistical model used in the analysis will be hierarchical regression. The calculation of the sample size used the following assumptions about the data analysis models. The dependent variables were related to driving performance or attention during automation. The independent variables were based on all the psychophysiological measures. To estimate the necessary sample size, small, medium and big effect sizes were used based on Cohen's f. The maximal number of independent variables included in the regression equation could be over seventy that included all the physiological factors, questionnaires, condition and time of the day. However, the problem of overfitting of the regression models makes it more reasonable to reduce the number of factors for one model and compare different models among each other during the later analysis. Because of that, the maximal number of the independent variables that was assumed for one model was ten. The sample size was calculated for multiple linear two-tailed regression in G*Power software. For small effect size, the recommended sample size was two hundred fifty-four, for a medium effect size one hundred and eight, and large effect size seventy-one. The project required a very long experimental process. Also, each participant was remunerated for their time and effort. Because of the financial and temporal restraints of the project, the assumed sample size was one hundred and two (fifty-one participants, each of them tested twice). This amount was in between the minimal sample size for large and medium effect and was the highest feasible amount that could meet the financial and temporal restraints of the project.

PARTICIPANTS

Sixty individuals were recruited for the experiment. They were healthy males and females with a full driving licence, who had not experienced motion sickness in the past. They were all informed about the procedure of the experiment and signed the informed consent.

All the participants were required to come to the laboratory twice, once for the night-time experiment and once for the day-time experiment. It was assumed that the second experimental session might be characterised by different driving performance due to the learning effect. Because of that, the sessions were scheduled in the way that half of the participants had a night session as their first experimental session, and the other half a day session as their first experimental session. That way, the learning effect would not confound the impact of the circadian phase. In the further description of the results, the term session was related to the variable indicating if the participant came to the lab for the first or for the second time.

Each session consisted of the same experimental scenarios, including two manual driving tasks. So, the task described as a first driving task, second session, would indicate the second time when participant came to the lab, and the first driving task of this experiment.

Eight participants had to terminate the experiment due to motion sickness. Out of the remaining fifty-two participants, three participated only in one experimental session, and forty-nine completed two full experimental sessions. Despite the attempts to recruit participants with an even distribution over genders and age groups, there was a particular bias towards male and young participants. The mean age of the participants who did not experience motion sickness was 28.33, and the median age was 26 years old. The demographics of the participants were presented in Figures 8.1, 8.2, and 8.3.

The psychophysiology of motion sickness was not a topic of this thesis; however, incomplete data from the participants who experienced motion sickness was analysed and published as a conference paper. The paper depicted some differences between the participants with and without motion sickness, mostly related to the breath, sleepiness, and the number of hours they have slept before the experiment (Kaduk et. al 2021.

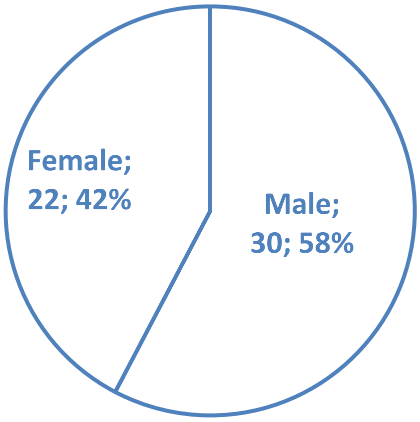

Figure 8.1: The gender distribution of the participants of the experiment.

As shown in figure 8.1, 58% of the participants were male and 42% were female.

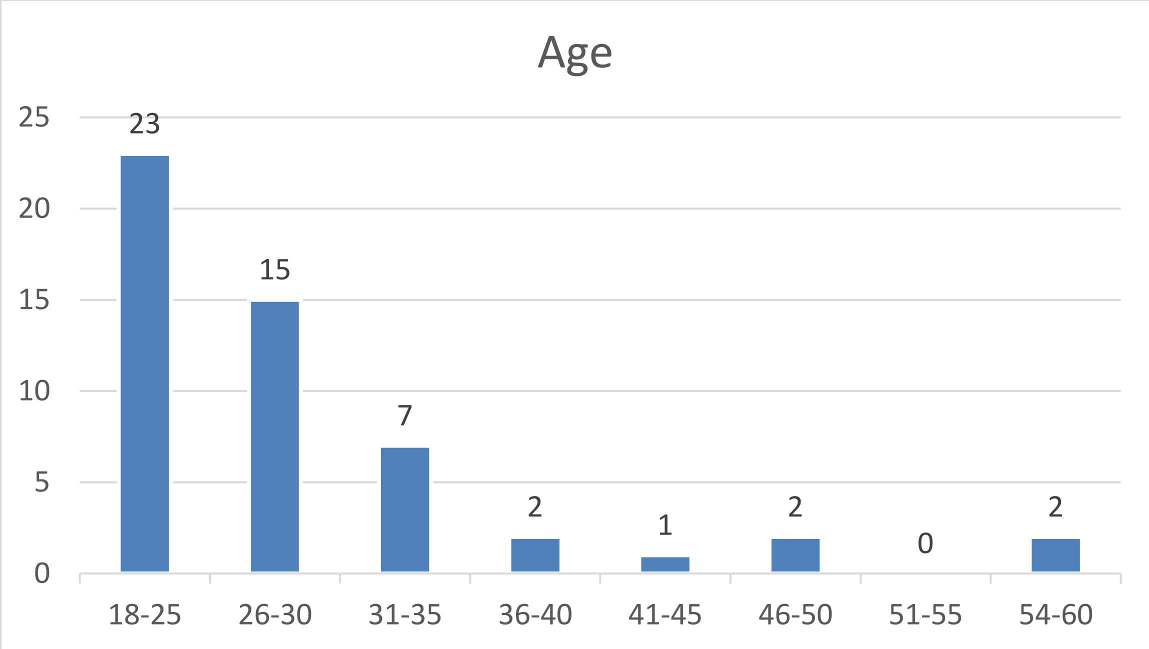

Figure 8.2: The age distribution of the participants of the experiment.

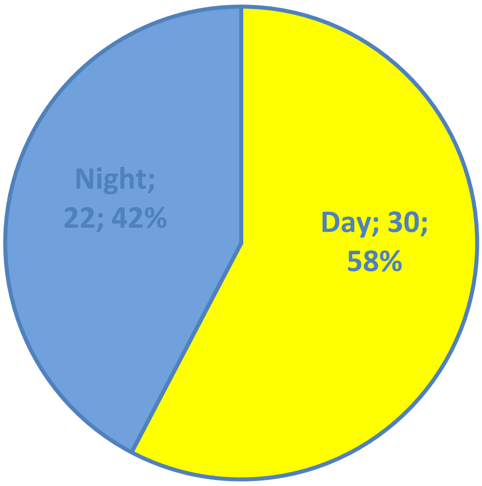

Figure 8.3: The number of participants who started the experiment from the day vs from the night session.

As shown in figure 8.3, 58% of the participants had a day-time session as their first session, and 42% had the night-time session as their first experimental session.

Several technical and participant related reasons led to the loss of some data recordings. Tables 8.1 present the number of recordings gathered with different measures, and table 8.2 the number of questionnaires collected and completed driving tests.

Table 8.1: Number of participants and measures collected with different methods.

	Number of all recordings	Number of Recordings without Motion Sickness	Number of Participants	Number of Night recordings	Number of Day Recording
EEG	47	47	27	23	24
ECG	108	101	52	50	51
Respiration	107	100	52	49	51
EMG	57	50	26	26	24
EOG	56	49	25	25	24
EDA	99	92	51	45	47

Table 8.2: Number of participants and measures collected with different questionnaires and driving test methods.

	KSS	Fatigue scale	TORS	NASA-TLX	Driving Test
M1	101	99	91	89	
T1					99
M2	101	98	90	90	
M3	99	99	90	88	
T2					100

8.2.2 ENVIRONMENT OF THE EXPERIMENT

The experiment was conducted in a controlled laboratory environment. Driving tasks were performed in the low-fidelity driving simulator. The driving simulator was placed inside the enclosed Faraday Cage (see chapter 7). The simulator comprised of the set of three screens placed in the way to simulate the surrounding of the participant, driver seat, steering wheel, and the set of pedals. The software of the driving simulator was STISIM 3. The set-up of the driving simulator can be seen in Figure 8.4. The term ‘low-fidelity driving simulator’ corresponds to the fact that the experiment used a driver seat, steering wheel, and pedals instead of the whole vehicle. It does not correspond to the quality of the simulation itself, realism or the software used. It means, therefore, low physical fidelity rather than low task fidelity (Roberts et al., 2020).

Figure 8.4: The set-up of the driving simulator inside the Faraday Cage.

Psychophysiological measurements were conducted with Enobio-20 EEG device, BioNomadix BioPac devices for EMG, EOG, EDA, ECG and respiration measures, Nellcor n-595 oximeter, Zoom H4N pro voice recorder. Saliva was collected with Sallimetrics cotton swabs. Questionnaires used were Karolinska Sleepiness Scale (KSS) (Akerstedt et al., 2014; Akerstedt & Gillberg, 1990), Samn-Perelli Fatigue Scale (Samn & Perelli, 1982), NASA-TLX mental workload scale (Hart & Staveland, 1988) and self-created take-over readiness scale (TORS) that can be seen in Appendix 3. The details of the measurements' choice can, experimental set-up, and laboratory construction can be found in Chapter 7.

8.2.3 EXPERIMENTAL PROCEDURES

Each of the participants participated in the same experimental session twice, once during the day-time circadian hours of high performance between 9 a.m. and 1 p.m. and once during the night, during the circadian low-performance hours between 10 p.m. and 2 a.m. Each of the experimental sessions had different durations due to the individual time of the task completion. The hours of the experiments provided above was then a given time range, but many of the experimental session finished earlier.

The experiment consisted of a challenging manual driving task, subsequent automated driving task and then manual driving task again. Such a configuration was aimed to simulate a driving sequence when a driver drives manually, then enters the automated mode and then is required to take-over the manual driving. Due to the high amount of psychophysiological measurements, take-over was not realistically simulated but was rather a break for the measurements collection between the automated mode and the second manual driving task. The take-over simulated in this experiment was a particular type of take-over when manual driving was resumed in a planned way when the vehicle left the area designated for automation rather than a take-over caused by a sudden emergency. This work did not analyse emergency take-over; however, some parts of it could be applied to such an event. This enabled investigation into the effect of automation on the subsequent manual driving as well as attention processes during the automated mode. The driving tasks were highly challenging to achieve the sensitive measurement of the small changes in the driving performance. The driving scenarios consisted of the sequence of the same driving scenes but presented in different order to reduce learning effects.

The EEG (results not included in this thesis), ECG, EOG, EMG, respiration and EDA were continuously measured during the whole scenario. Still, only some parts of the recording were analysed later. Voice, pulse, blood oxygenation and saliva measures were collected only at some points of the experiment. The experimental procedure is presented in Figure 8.5.

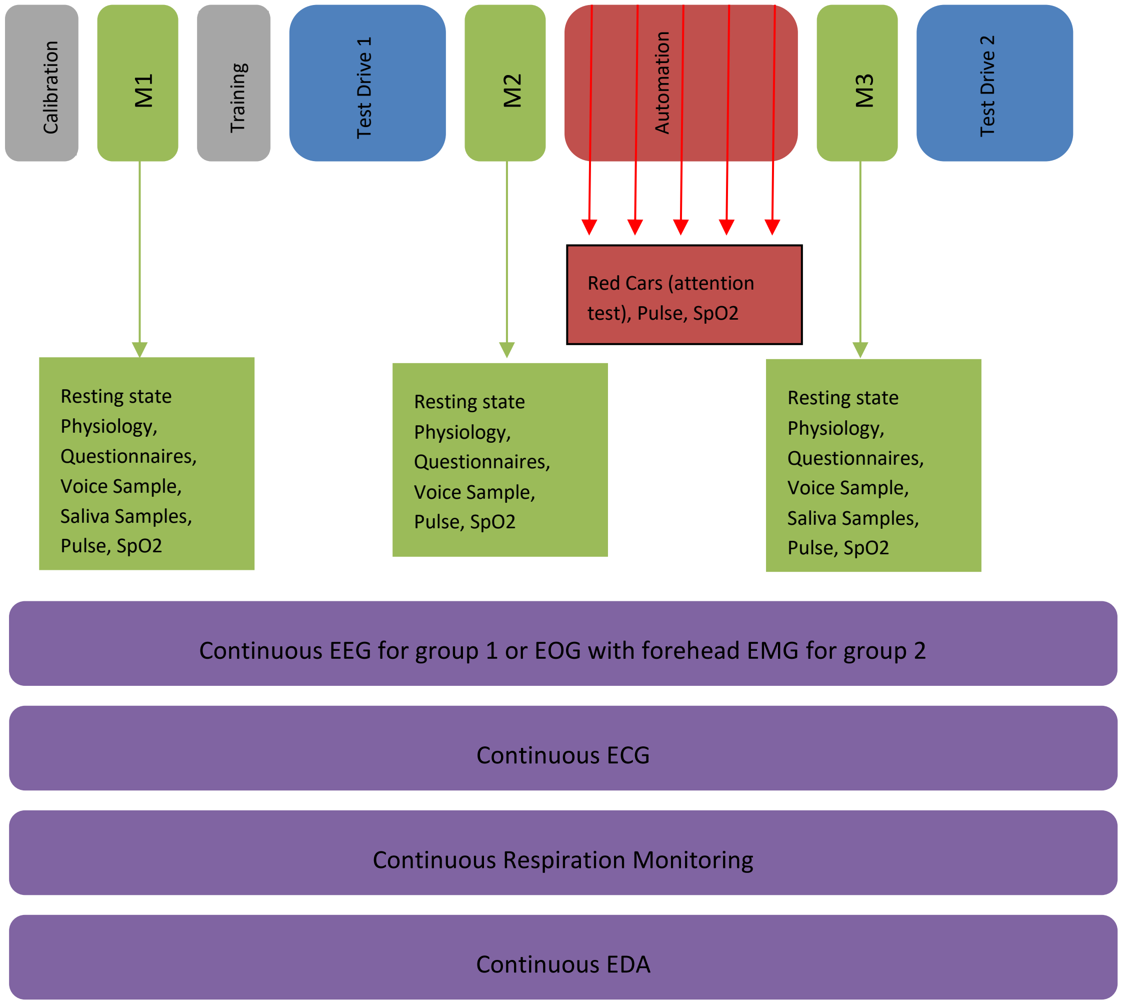

Figure 8.5: Graphical representation of the stages of the experiment. Grey boxes represent additional stages, green stages with a discrete collection of psychophysiological measurements, blue stages with the manual driving measures, red with automated driving measures, and purple continuous physiological measurements.

The calibration of the devices was in the first phase of the experiment when participants were asked to perform different activities such as blinking or clenching the teeth to establish sample signals for noise removal during the data pre-processing.

During Measurement 1, Measurement 2, and Measurement 3 participants were asked to sit still for two minutes with eyes open to collect the resting measures of EEG, ECG, EOG, EMG, EDA, respiration, pulse, and blood oxygenation. Subsequently, they were asked to read the sentence ‘good bed is more positive than a hot toe' that contains many of the voiced and unvoiced consonants for the acoustic voice analysis. The two samples of the saliva were collected during Measurement 1 and Measurement 3 (four in sum) for the cortisol and alpha-amylase analysis. Set of questionnaires: KSS, Fatigue scale, NASA-TLX and TORS were given to the participants at each measurement.

Training Drive was a concise driving scenario that participants undertook to become familiar with the operation of the driving simulator.

Test Drive 1 and Test Drive 2 (later called Test 1 and Test 2 or T1 and T2) were very challenging driving scenarios containing several scenes with the driving challenges or unexpected change that drivers had to react to. The order of the scenes was randomized to reduce the learning effect on the scores. Each scene contained a particular type of driving challenge: driving through a city with traffic lights and pedestrian crossings, driving through a village, long period of highway with cars suddenly changing lane, fog, driving up the hill through the narrow, steep road with low visibility, driving down the hill through the slippery road and going into the slide, and construction on the road and the lorry suddenly blocking the middle of the road. Test Drives were the sensitive measures of the participants' driving performance. Before each driving task participants were asked to drive as well and as accordingly to the rules as they can.

Automation was the period when the car was driving itself through the area of Southampton. Participants were asked to press a key whenever they have seen a Red Car as a measure of attention. There were five Red Cars in the thirty-four minutes of the automated scenario. The number of Red Cars detected was a measure of attention. Before the automated task participants were asked to stay as attentive as possible and monitor the progress of the vehicle.

Participants had two breaks after Measurement 1 and Measurement 3 when they were offered water and snacks. It was an ethical requirement due to the long duration of the experiment. Except for the snack breaks, each stage of the experiment was performed in sequence. Participants were allowed to ask for a comfort break at any time. As so, in some cases, there were longer intervals between the stages of the experiment. It could create a confounding influence; however, due to the long duration and tiring circumstances, it would be unethical to enforce a strict schedule onto participants. Also, the duration of the whole experiment differed for every participant because the length of the driving tasks depended on the individual driving speed as well as differences in the time it took to complete the questionnaires. In sum, the experiment lasted between two and a half and four hours.

8.2.4 MEASURES USED IN THE PILOT STUDY AND THE MAIN EXPERIMENT

Due to the technical malfunction, EDA was not used in the pilot experiment, but only in the main experiment after being repaired by the company. Eye-tracker was completely excluded after the pilot experiment because of the high frequency of software and hardware failures, interference with electrodes and headache caused after a long period of wearing it together with the forehead electrodes. As a result, the following devices were used in the pilot experiment ECG, respiration belt, EMG and EOG BioNomadix models with MP150 module by BioPac. Voice was recorded with Zoom H4N pro recorder, Sennheiser phantom power adapter MZA 900 P, and Sennheiser head-mounted microphone HSP 4. Blood oxygenation and pulse were measured with a Nellcor n-595 oximeter. Saliva was collected with Salimetrics oral cotton swabs and frozen at -26 Celsius degrees; however, analysis of the saliva was conducted cumulatively after all the samples were gathered. Driving task and driving performance measurements were conducted with STISIM 3 driving simulator with the steering wheel, driver's seat and pedals. There was also an additional attention task during the automated mode. Participants were asked to press a button every time they saw a red car. The number of red cars detected was treated as an indicator of attention during automation. Questionnaires used for subjective state measurements were Karolinska Sleepiness Scale (KSS) for drowsiness/sleepiness assessment (Akerstedt et al., 2014; Akerstedt & Gillberg, 1990), NASA-TLX for the mental workload assessment (Hart & Staveland, 1988), and Samn-Perelli scale for the fatigue assessment (Samn & Perelli, 1982). There was also a scale created for this experiment to assess the subjective readiness of the participant for the driving take-over. The scale called TORS (Take-Over-Readiness-Scale) can be seen in Appendix 3. For the main experiment, the following methods were used: EEG, EMG, EOG, ECG, respiration, EDA, blood oxygenation, pulse, voice recording, saliva analysis, KSS questionnaire, Samn-Perelli Fatigue Scale, NASA-TLX questionnaire, TORS questionnaire, driving performance and attention during automation. It is important to note that even though the EEG signal was gathered during the main experiment it is not a subject of this thesis due to the high temporal demand of the data pre-processing. It is going to be pre-processed an analysed in future.

Driving performance was calculated with the indicators provided by STISIM 3: total number of the off-road accidents, total number of collisions, total number of pedestrian hits, total number of speed exceedances, the total number of speeding tickets, total number of traffic light tickets, total number of stop signs missed, total number of centreline crossing, the total number of road edge excursion, the standard deviation of the lane position, standard deviation of the steering wheel rate, the standard deviation of the vehicle heading angle, the standard deviation of the longitudinal speed and acceleration of the vehicle, and a variable summarizing general driving performance. Higher variables indicated worse driving performance.

Attention during the automated mode was calculated with the number of red cars detected during the automated scenario. The number of red cars in the thirty-four minutes simulation was five. A low number was aimed to challenge sustained attention.

Scores from KSS and Samn-Perelli scale were calculated as a straightforward indicator of the subjecting sleepiness and fatigue. Higher scores indicated a higher level of sleepiness or fatigue with nine points as a maximum in KSS and seven points as a maximum in Samn-Perelli Scale (Akerstedt et al., 2014; Akerstedt & Gillberg, 1990; Samn & Perelli, 1982). Scores from scales of NASA-TLX were calculated with the weighting in the way recommended in the manual (Hart & Staveland, 1988). TORS was measured each time and a score indicated by the participant was expressed in the form of a percentage, a higher percentage representing a lower level of readiness to take over manual driving. The questionnaires can be found in Appendix 3.

8.2.5 EXPERIMENTAL SET-UP

While careful laboratory construction allows reduction of the number of potential noise sources, it is impossible to eliminate all of them. Many measures are susceptible to body movements or muscle spasms (Chong et al., 2014; Taylor et al., 2015). Because of this, there is a need for a careful devices set-up and some preparations. Also, before each experimental session signal from the devices was checked through the process of calibration. Participants were asked to perform some actions or movements and recordings were checked in terms of proper reaction. For example, participants were asked to blink and move their eyes, and the EOG recording was checked for the movement signal. If the devices did not display proper signal, they were repaired until it was obtained. The following sections described additional activities undertaken to improve the signal from each of the devices. The impedance of the ECG, EMG, and EOG electrodes was checked after application with electrodes impedance checker and electrodes were repaired if the impedance was too high. The experimental set-up of the devices can be seen in figures 8.6

and 8.7.

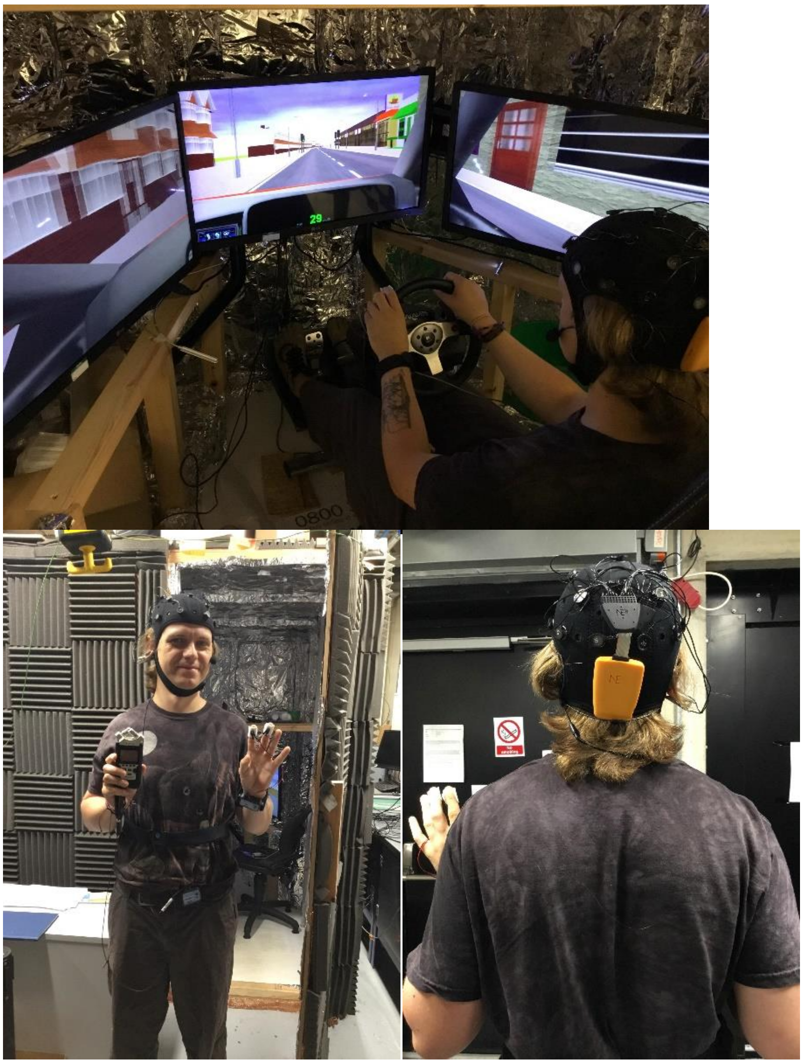

Figure 8.6: Experimental set-up of the participant allocated to group 1 with EEG measurements.

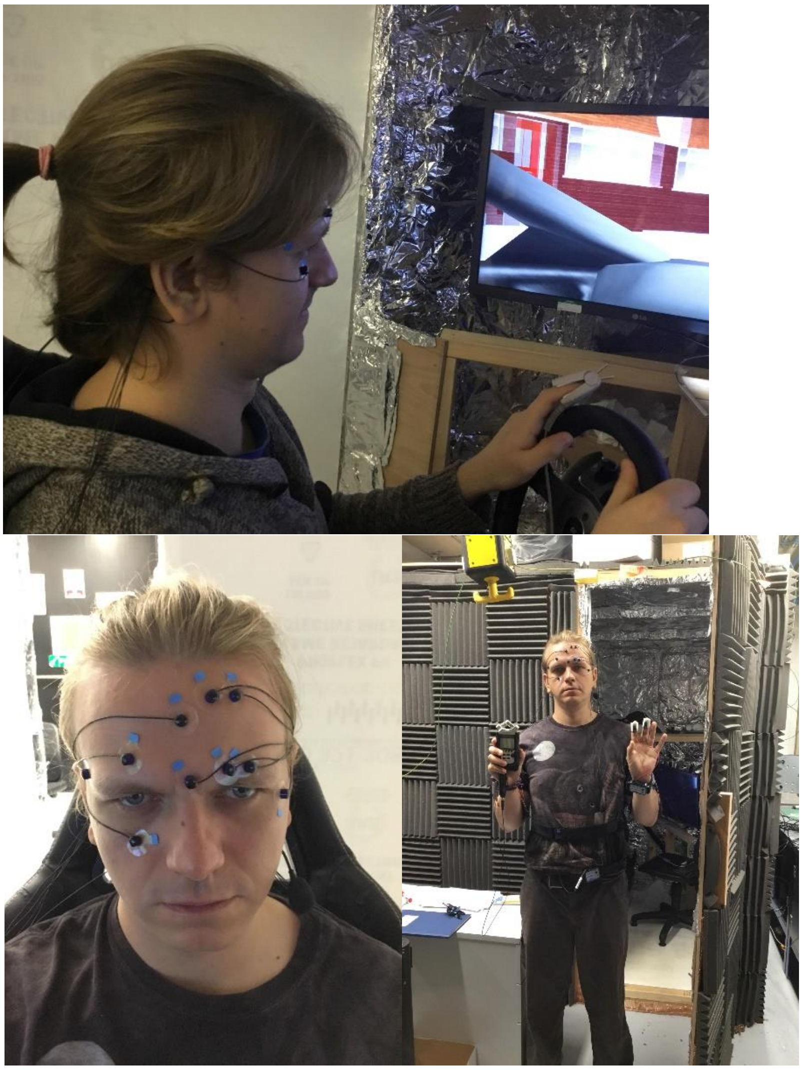

Figure 8.7: Experimental set-up of the participant allocated to group 2 with EOG and forehead EMG measurements.

ELECTROENCEPHALOGRAPHY

EEG was collected with Enobio20 device, with geltrodes and foretrodes by Neurolectrics as electrodes. The scalp of the participants was thoroughly prepared for the electrodes through scrubbing with a NuPrep gel. The electrolytic Signa Gel was used to enhance electrical conductivity. The pattern of movements was recorded during the calibration. Also, the impedance of the electrodes was monitored continuously. The skin was additionally cleaned and scrubbed in case of an impedance increase.

ELECTROOCULOGRAPHY

EOG was measured separately from EEG. The signal was recorded with BioNomadix EOG 2 amplifier and MP150 BioPac module. Electrodes used were BioPac EL254 Ag-AgCl EMG electrodes. The skin around the eyes was prepared with NuPrep gel, and the conductivity was increase with SignaGel. The quality of the signal was first checked during calibration and then monitored during the whole experiment. In case of signal quality decrease electrodes were disconnected, cleaned and applied again after renewed skin preparation.

ELECTROMYOGRAPHY

EMG was measured separately from EEG as the cap was displacing forehead electrodes. The signal was recorded with BioNomadix EMG 2 amplifier and MP150 BioPac module. Electrodes used were BioPac EL254 Ag-AgCl EMG electrodes. The skin was prepared with a NuPrep gel, and conductivity was increased with a SignalGel. During calibration, participants were asked to perform several forehead movements to check the quality of the signal. Later, the quality was monitored continuously during the experiment. In case of the quality decrease, electrodes were removed, cleaned, and applied again after renewed skin preparation.

ELECTROCARDIOGRAPHY

The ECG signal was collected with a BioNomadix ECG amplifier and MP150 module. In this experiment, the skin of the participants was prepared with the NuPrep Gel. The electrodes used were disposable electrode stickers manufactured by BioPac, and they already contained electrolytic gel. ECG has a very distinctive signal with QRS peaks. This signal was monitored continuously during the experiment. In the case of signal distortion, electrodes were removed from the chest of the participants, the skin was prepared again, and the new electrode stickers were applied.

RESPIRATION

The respiration signal was collected with a respiration belt by BioPac connected with an ECG amplifier, and MP150 module. The respiration signal was monitored continuously during the experiment, and in case of significant amplitude reduction, participants were asked to breathe deeply. If the signal did not display deep breathe, the belt placement was improved as it sometimes slipped down the participants' abdomen.

ELECTRODERMAL RESPONSE

EDA was recorded with a BioNomadix PPGED amplifier and MP150 module. In this experiment, EDA electrodes were placed on the index and middle finger of the participants' non-dominant hand. Participants were asked to wash their hands only with water just before the experiment and refrain from using any hand creams. Their fingers were cleaned with a dry cotton swab before application of the electrodes. The electrodes used were BioPac disposable EL507 electrodes, which were enhanced with an additional 102 BioPac Gel that induces signal conductivity without confounding it.

ACOUSTIC SPEECH ANALYSIS

In this experiment microphone used was head-mounted microphone HSP 4, with Sennheiser phantom power adapter MZA 900 P, and Sennheiser Zoom H4N pro voice recorder. The choice of the tools was based on the professional advice of the specialized company.

OXIMETRY

In this experiment, the light inside the laboratory was kept on a constant stable level. The oximetry sensor was placed on the ring finger of the non-dominant hand of the participant. The finger was previously cleaned with a dry cotton pad.

SALIVA COLLECTION

Cortisol and alpha-amylase saliva analysis are analytical procedures on human or animal saliva. To avoid disturbances in the analysis participants cannot consume food and beverages other than water for sixty minutes before the saliva collection. Consumption of medications, drugs and nicotine should be reported, as well as vigorous exercises before the collection. In the case of cortisol and alpha-amylase saliva can be collected with a cotton swab or passive drool. The cotton swab should be placed in the bottom area between the teeth and the cheek. After the collection saliva must be frozen immediately at a temperature lower than -20 Celsius degrees. If stored at a temperature of over -80 Celsius degrees samples must be analysed within a period of four months (Salimetrics & SalivaBio, 2011).

In this experiment, participants were asked to refrain from eating for an hour before the experiment. They were also requested only to drink water. During the experiment, they were offered snacks just after each of the saliva collections, so there was at least an hour break between the snack and following saliva collection. They received cotton swabs and were asked to keep them in between their upper chick and the teeth for two minutes. After collection, cotton swabs were placed in the labelled tubes and frozen at -26 degrees. The samples were shipped to the Anglia Ruskin BioLab in dry ice packaging no longer than three months after collection. At the laboratory, they were stored at -80 degrees until the analysis.

8.2.6 DATA ANALYSIS

The data analysis consisted of three major steps: data pre-processing, signal analysis and statistical analysis.

DATA PRE-PROCESSING

Pre-processing of the physiological data is the process of noise rejection to keep the only physiological signal for further signal analysis. The most commonly used methods are filters that reject frequencies that are not possible for the particular physiological function or manual artefact rejection based on the visual inspection of the signal (Berntson et al., 1997; Cacioppo et al., 2007; Delorme & Makeig, 2004).

ECG was pre-processed in AcqKnowledge software. First, it was filtered with a bandpass filter with a 0.05 Hz lower border and 35 Hz higher border with 8000 coefficients. After that, a peak identification function was used, peaks that were not identified or inter-beat periods that were identified as peaks were supposed to be manually corrected. However, the data was of high quality and there was only one recording that needed manual correction (Berntson et al., 1990, 1997; Berntson & Stowell, 1998).

Respiration was pre-processed in AcqKnowledge software. First, it was resampled into 50 samples per second rate to reduce the computational weight of the file. After that, it was filtered with a band-pass filter with 0.01 Hz lower border and 1 Hz higher border with 4000 coefficients (Kim et al., 2007; Lanatà et al., 2009; G. Liu et al., 2013).

EMG was pre-processed in AcqKnowledge software. The data was filtered with a 5 Hz high pass filter and visually inspected. In the case of noisy data, parts of the recording were rejected from the analysis (De Luca et al., 2010; Van Boxtel, 2001).

EOG was pre-processed in Acqknolwedge software. It was filtered with a band-pass filter with 0.1 Hz lower border and 20 Hz higher border (Banerjee et al., 2013).

EDA was pre-processed in Acqknolwedge software. It was first resampled into 50 samples per second rate to reduce the computational weight of the file. After that, it was smoothed with a median smoothing and filtered with a low pass 1 Hz filter (Braithwaite et al., 2013).

In the case of voice recordings, there was no need for artefact rejection; however, the recordings were cut into small pieces containing the only sentence of interest in the PRAAT software (Boersma & Van Heuven, 2001).

Oxymetry did not require pre-processing as the Nellcor device was giving pre-processed and analysed output.

SIGNAL ANALYSIS

Signal analysis is a branch of knowledge and techniques of processing pseudo-continuous signals into factors like time, frequency or power (Allen & Mills, 2004). The methods of signal analysis chosen for this work were based on the literature review on the driver state monitoring and technical manuals.

ECG was analysed with AcqKnowledge 5 software. Heart rate was calculated with a 'find rate' function. Heart rate variability was calculated both in a statistical and frequency manner (Brookhuis & de Waard, 2010; Fairclough & Houston, 2004).

The respiration signal was analysed with AcqKnolwedge 5 software. The 'Find rate' function was used to calculate breathing frequency, and the standard deviation of the original signal was calculated to estimate the mean breathing depth (Rodríguez-Ibáñez et al., 2011; Schreinicke et al., 1990).

EMG signal was analysed with AcqKnowledge 5 with 30 seconds epoch frequency power analysis (Van Boxtel, 2001).

EOG was analysed to estimated blinking rate, mean blinking duration, PERCLOS, rate of horizontal eye movements and mean duration of horizontal eye movements (Borghini et al., 2014; Rodríguez-Ibáñez et al., 2011). The signal was collected with Acqknolwedge 5; however, this software does not provide a function for EOG factors calculation. Signal was analysed using Matlab R2020a with a self-written code using two columns with voltage as input data. Code can be found on the author's GitHub account (https://github.com/SylwiaKaduk).

EDA was analysed in Acqknolwedge 5. Amplitude and frequency were calculated for focus areas to obtain parameters of the skin conductance level (Braithwaite et al., 2013).

Voice was analysed using PRAAT software to estimate the following voice properties mean pitch, maximum pitch, minimum pitch, voice frequency range, the standard deviation of the pitch, number of pulses, number of periods, the fraction of locally unvoiced frames, number of voice breaks, degree of voice breaks, jitter, shimmer, mean autocorrelation, mean noise to harmonics ratio, mean harmonics to noise ration, mean speech intensity, maximum speech intensity, and minimum speech intensity (Boersma & Van Heuven, 2001).

Oximetry did not require signal analysis as the Nellcor device was giving pre-processed and analysed output. Blood oxygenation and pulse were averaged over the two minutes of

Measurement 1, Measurement 2, and Measurement 3. It was also recorded at the moment of each red car appearance in the automated mode.

STATISTICAL DATA ANALYSIS

Each of the physiological variables was tested for normality with the Kolmogorov-Smirnov test. All tests indicated non-normal distributions of the variables, so the statistical tests used in the analyses were non-parametric ones.

To analyse a time-course of psychological and physiological variables were centralized. Centralization was meant to allow both within and between subjects' analysis. Kruskal-Wallis test and multiple comparisons tests with Bonferroni corrections were used to compare psychophysiological measures at Measurement 1, Measurement 2, and Measurement 3. Driving performance was not centralized as there were only two measurement points during the experimental session and centralization would not allow seeing to the magnitude of the change between the driving tests. Because of that, driving test 1 and test 2 were compared with Wilcoxon signed ranks test. The statistical analysis and data processing were conducted in Matlab R2020a and Excel.

8.3 THE TIME COURSE OF SEMI-AUTOMATED DRIVING-EXPERIMENTAL RESULTS

8.3.1 MANUAL DRIVING PERFORMANCE

The Wilcoxon signed ranks test was used to compare different factors of the driving performance during the first (T1) and second (T2) manual driving tasks. Only the second experimental sessions were analysed to avoid the confounding effect of learning on the difference between T1 and T2. The standard deviation of lane position, a standard deviation of steering wheel angle, a standard deviation of the vehicle heading angle, a standard deviation of vehicle's longitudinal speed, and general driving performance significantly differed between the driving tasks. All the factors indicated worse driving performance after the automated phase. The exact values of the factors that significantly differed between T1 and T2 were presented in Table 8.4. The differences were shown in Figures 8.8 and 8.9.

Table 8.3: Factors affecting driving performance that significantly differed between T1 and T2. A higher value of the factors indicates worse driving performance.

	Z value	P value	Mean before automated phase	Mean after automated phase	SD for T1	SD for T2	SD for T1 and T2
Standard Deviation of Lane Position	-2.55	<.05	10.11	12.18	6.54	8.74	7.32
Standard Deviation of Steering Wheel Angle	-4.85	<.05	255.73	287.11	37.85	43.35	41.20
Standard Deviation of Vehicle Heading Angle	-3.41	<.05	123.30	135.21	20.50	5.48	14.29
Standard Deviation of Vehicle Longitudinal Speed	-3.25	<.05	27.17	30.03	5.92	5.54	5.67
General Driving Performance (higher the value, lower the performance)	-2.84	<.05	1172.76	1254.88	181.95	166.74	169.98

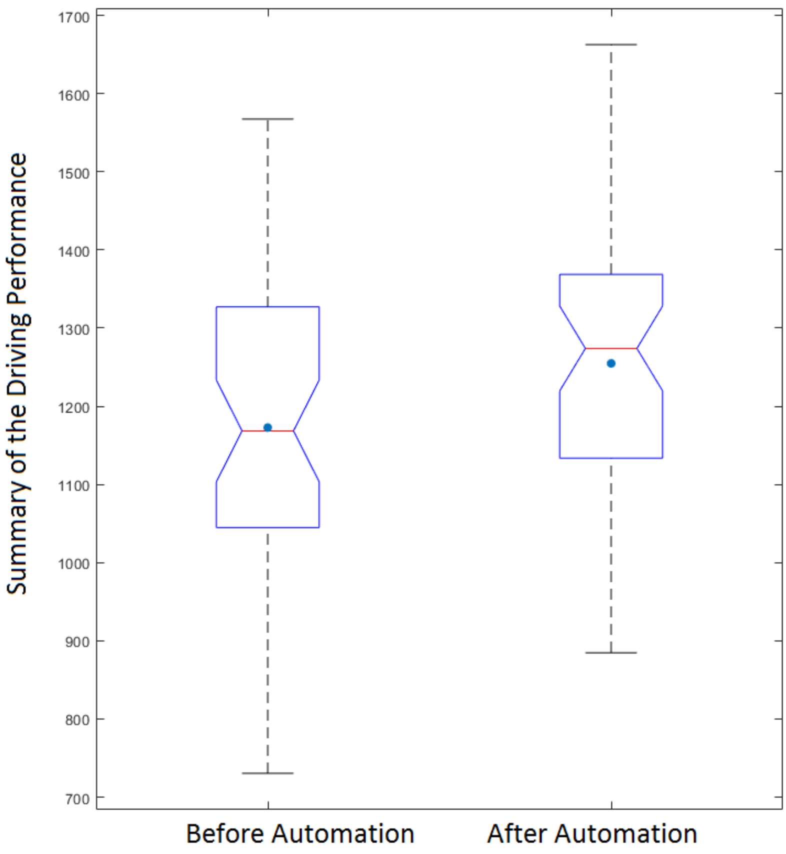

Figure 8.8: Comparison of general driving performance before and after automation. A dot represents the mean value. A higher value of the factors represents worse driving performance.

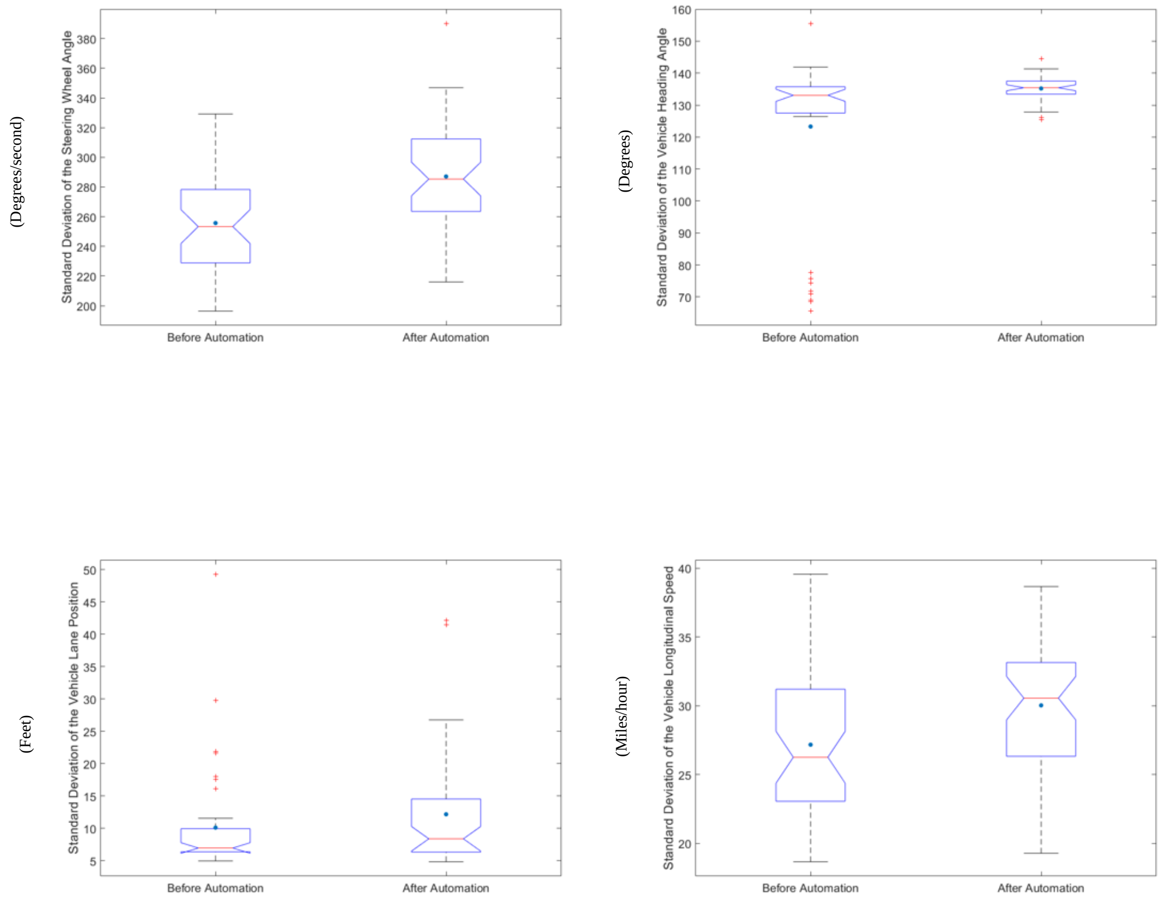

Figure 8.9: Comparison of different driving performance factors before and after automation. A dot represented the mean value. The plots represent following factors of driving performance: left-up: standard deviation of the steering wheel angle, right-up: standard deviation of the vehicle heading angle, left-down: standard deviation of the vehicle lane position, right-down: standard deviation of the vehicle longitudinal speed.

As an additional analysis, the same comparison was conducted separately for the day and night experiments. For both circadian phases, there was a decrease in driving performance after the automated phase; however, during the night experiment, there were more factors that significantly deteriorated as presented in tables 8.5, 8.6, and figure 8.10.

Table 8.4: Factors in driving performance that significantly differed between T1 (before automation) and T2 (after automation) for the day-time experimental session.

	Z value	P value	Mean before automation	Mean after automation	SD for T1	SD for T2	SD for T1 and T2
Standard Deviation of Steering Wheel Angle	-2.97	<.05	258.02	292.95	45.03	51.53	48.59
Standard Deviation of Vehicle Heading Angle	-3.04	<.05	122.57	136.44	17.63	6.98	12.69
Standard Deviation of Vehicle Longitudinal Speed	-1.93	<.05	27.99	30.23	6.25	5.25	5.73

Table 8.5: Factors in driving performance that significantly differed between T1 (before automation) and T2 (after automation) for the night-time experimental session.

	Z value	P value	Mean before automation	Mean after automation	SD for T1	SD for T2	SD for T1 and T2
Standard Deviation of Lane Position	-2.53	<.05	8.76	12.68	5.88	8.49	6.91
Standard Deviation of Steering Wheel Angle	-3.62	<.05	255.09	281.83	29.17	33.32	32.04
Standard Deviation of Vehicle Heading Angle	-1.89	<.05	124.06	134.38	23.00	3.33	15.71
Standard Deviation of Vehicle Longitudinal Speed	-2.76	<.05	26.52	29.82	5.62	5.87	5.62
General Driving Performance (higher the result, lower the performance)	-2.27	<.05	1155.18	1238.59	170.40	142.58	151.91

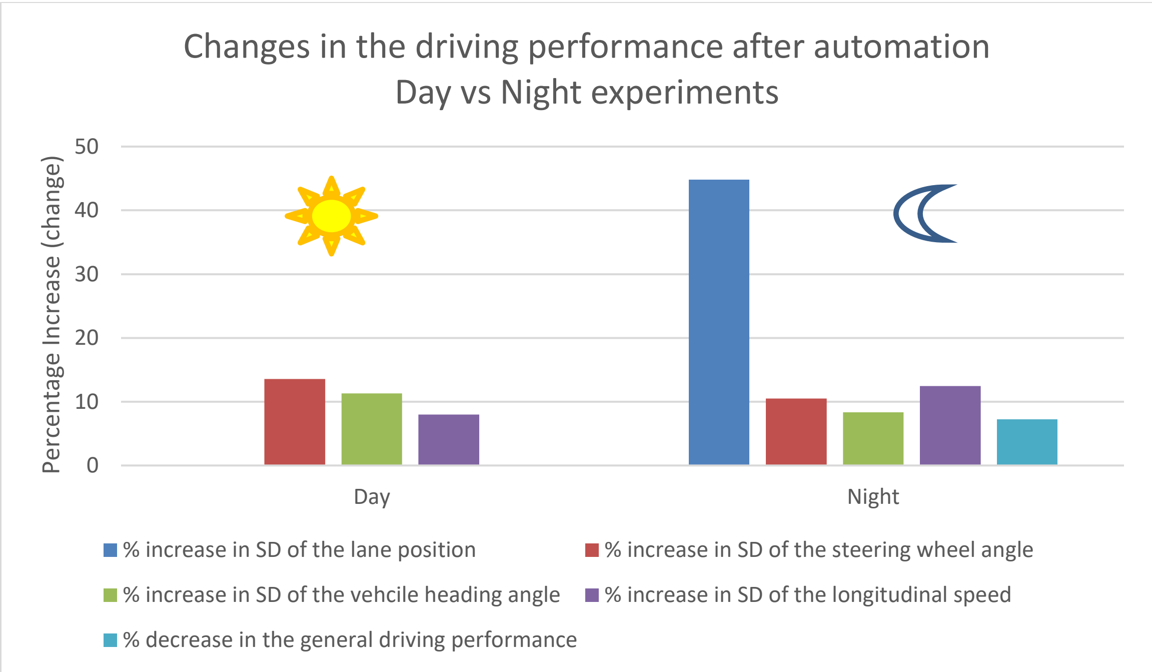

Figure 8.10: % change in the driving performance between T1 and T2, day and night comparison. The higher increase in scores represents a larger decrease in driving performance.

8.3.2 QUESTIONNAIRES

The Kruskal-Wallis test with multiple comparisons was used to test differences between questionnaires' scores collected at M1 (baseline level), M2 (after manual driving), and M3 (after automation). Before conducting the statistical tests, questionnaires' scores were centralised to keep within-participant information during the between-participant analysis. Centralisation was applied to each participant and each session separately. Table 8.7 presented the p values of the multiple comparisons tests and mean values of the questionnaires' scores for the different measurement points.

Table 8.6: p values of the Kruskal-Wallis multiple comparisons with Bonferroni correction between questionnaires collected at M1 (baseline measurement), M2 (measurement after manual driving), and M3 (measurement after automation) with mean values for each measurement point.

	Chi^2	P value M1 vs M2	P value M1 vs M3	P value M2 vs M3	M1-baseline mean	M2-after manual mean	M3-after aut. mean	SD for M1	SD for M2	SD for M3	SD for M1, M2, and M3
KSS	172.75	NS	<.05	<.05	4.01	4.19	6.50	1.95	2.09	1.99	2.16
TORS (higher the score, lower the readiness to take-over)	139.96	<.05	<.05	<.05	1.96	2.38	3.43	0.77	1.08	1.31	1.14
Fatigue	184.04	<.05	<.05	<.05	2.57	2.91	4.15	1.27	1.29	1.27	1.37
NASA-TLX mental demand scale	89.29	<.05	<.05	NS	8.91	32.42	26.51	14.91	24.49	28.29	23.66
NASA-TLX physical demand scale	15.85	<.05	NS	<.05	6.10	9.87	7.57	12.32	12.36	14.18	12.20
NASA-TLX temporal demand scale	18.25	<.05	<.05	NS	6.19	12.78	11.90	11.45	16.97	19.33	15.26
NASA-TLX performance scale	64.22	<.05	<.05	<.05	9.03	22.78	12.89	15.52	19.94	12.79	15.93
NASA-TLX effort scale	124.17	<.05	<.05	NS	4.63	32.51	35.57	9.69	24.28	26.20	22.99
NASA-TLX frustration scale	122.41	<.05	<.05	<.05	5.54	39.80	31.81	11.26	26.08	27.55	24.35
NASA-TLX general mental workload	162.23	<.05	<.05	<.05	2.60	9.82	8.38	2.90	5.19	5.05	4.80

Sleepiness significantly increased after the automated phase, but not after manual driving. Participants felt less ready to take-over manual driving after they went through the manual driving task, and then even less after the automated phase. Fatigue increased after manual driving and then almost doubled after the automated phase. Mental workload showed a different tendency. The majority of scales from NASA-TLX showed a trend to increase after manual driving and decrease after automation; however, not all the changes were significant. Only the Effort scale scores increased after automation, but the increase was not significant. Figures 8.11, 8.12, 8.13, and 8.14 showed changes in sleepiness, readiness to take-over, fatigue, and general NASA-TLX scores over the time-course of the experimental sessions.

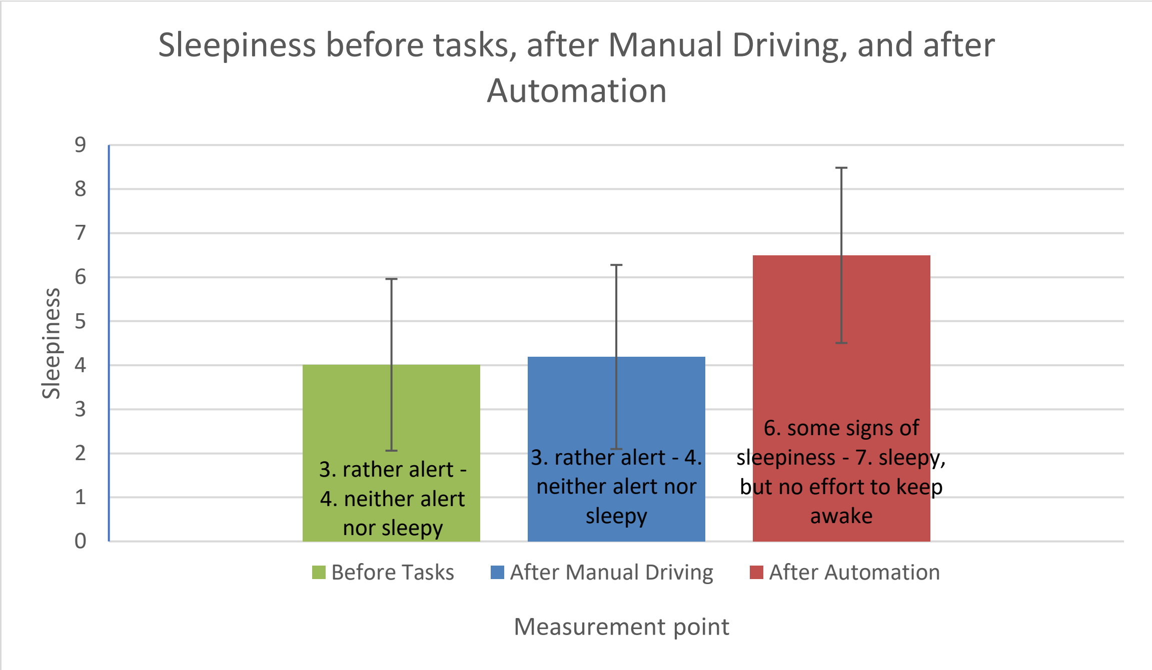

Figure 8.11: Mean Values with standard deviations of the Karolinska sleepiness scale during the experimental session. The change between M1 (baseline measurement) and M2 (measurement after manual driving) was not significant.

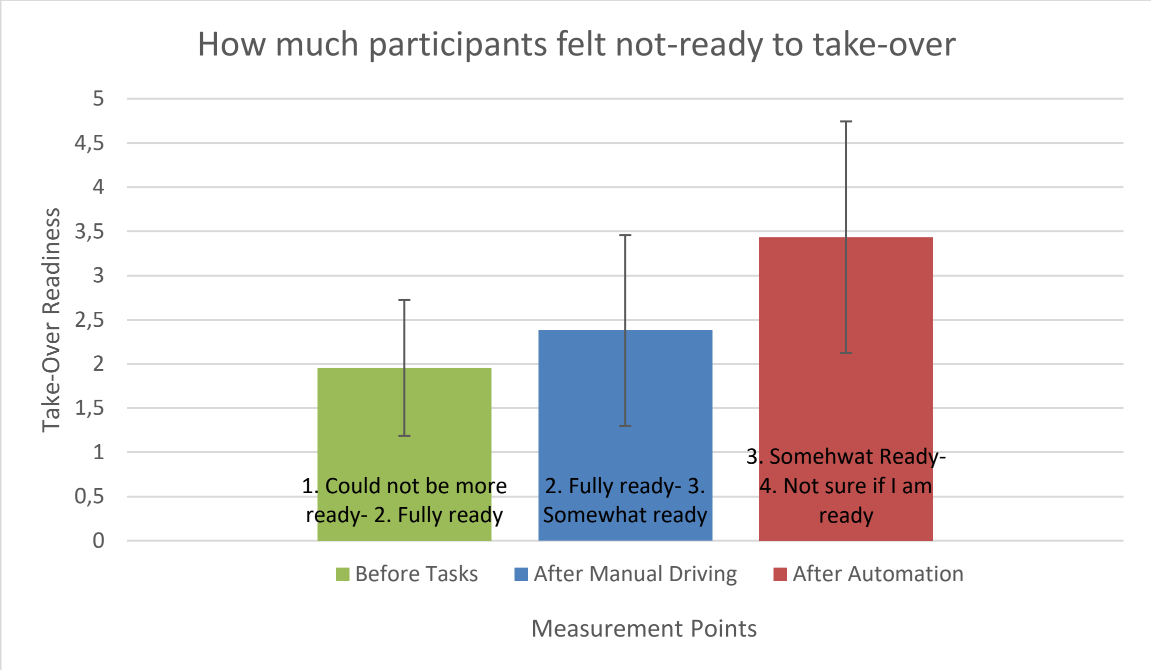

Figure 8.12: Mean values the Take-Over Readiness Scale with standard deviation during the experimental sessions. Higher results on the scale indicate lower readiness to take-over manual driving.

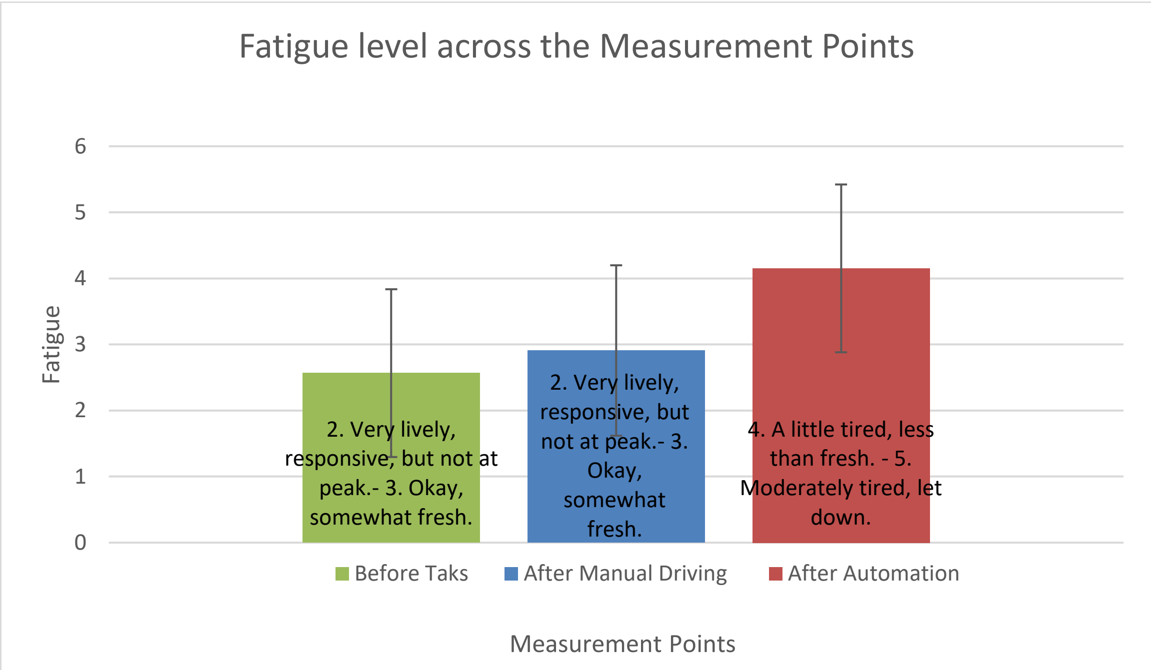

Figure 8.13: Mean Values of the Samn-Perelli Fatigue Scale scores with standard deviation during the experimental sessions.

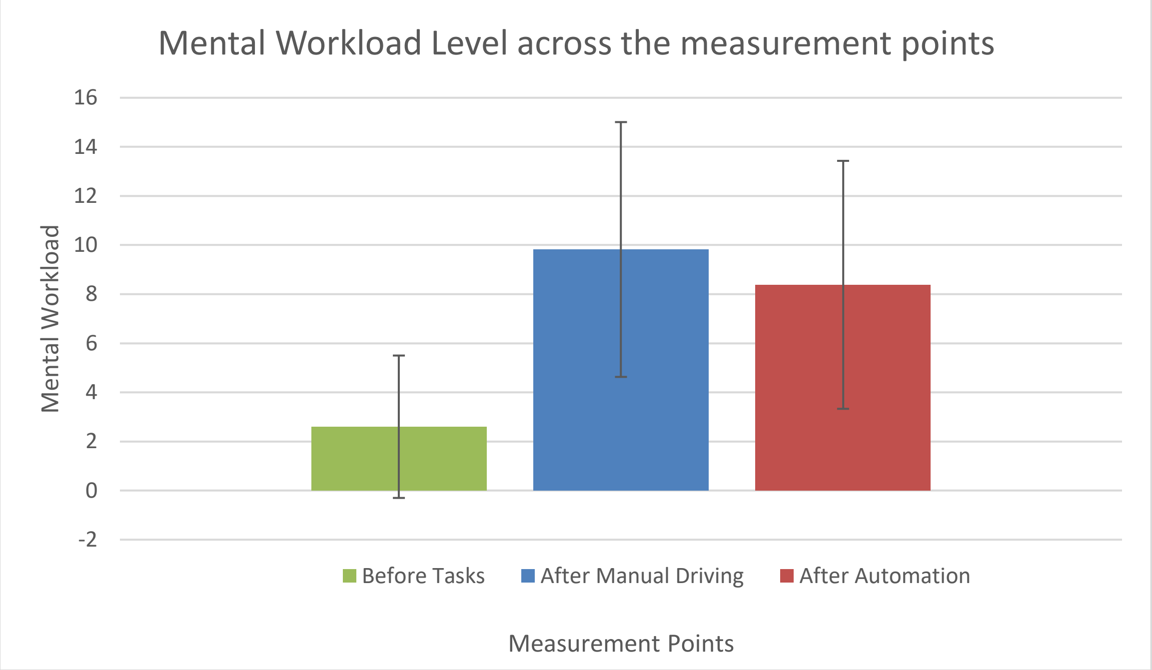

Figure 8.14: Mean Values of the general mental workload NASA-TLX scores with standard deviation during the experimental sessions.

The driving performance tended to decrease after using the automated mode. Sleepiness tended to increase as well as fatigue. Participants tended to feel less ready to take-over manual control of the car after using the automated mode in comparison to previous measurements. They also felt less ready to take-over manual control after first manual driving in comparison to the baseline measurement. It could be hypothesized that it was a matter of how challenging the driving task was or of the driver fatigue. At the same time, the mental workload showed a decreasing trend after activating the automated mode. To investigate whether the changes in the subjective state could address driving performance decreases, rho-Spearman correlations were calculated between all the driving performance factors and all the questionnaires' results. The majority of correlations were non-significant, and the significant correlations were low. After the Bonferroni correction only the correlation between overall NASA-TLX scores and standard deviation of the vehicle heading angle was significant ($rs(197)=0.34$, $p<.05$).

8.3.3 ELECTROCARDIOGRAPHY

To investigate the changes in ECG over the experimental session, the Kruskal-Wallis test was conducted to compare each of the ECG variables during M1, M2 and M3. The comparison was conducted on the centralised values of the variables to combine both within and between subjects' information. In the case of significant results, multiple comparisons test was conducted for further investigation. The results described below included only ECG variables that significantly differed between M1, M2, and M3.

Heart rate significantly differed between the measurement points. A multiple comparisons test showed that heart rate during baseline measurement was significantly higher than heart rate after manual driving, and heart rate after automation. But the difference in heart rate after manual driving and automation was insignificant.

Power in low heart rate variability frequencies significantly differed between measurement points. The multiple comparisons test showed that change after manual driving was insignificant, but there was a significant increase in power after automation.

Power in high heart rate variability frequencies significantly differed between measurement points. The multiple comparisons test showed a significant increase after manual driving but not after automation.

Power in heart rate variability sympathetic tone and vagal tone significantly differed between measurement points. The multiple comparisons test showed a significant increase of the sympathetic tone and a decrease of the vagal tone between the measurement of baseline level and measurement after automation but not between baseline level and after manual driving or after manual driving and after automation.

RSA in heart rate variability significantly differed between measurement points. The multiple comparisons test showed a significant increase between baseline measurement and measurement after automation but not between baseline level and after manual driving or after manual driving and after automation.

PNN50% in heart rate variability significantly differed between measurement points. The multiple comparisons test showed a significant increase between baseline measurement and measurement after manual driving and between baseline measurement and measurement after automation but not between after manual driving and after automation.

The detailed results of multiple comparison tests were presented in the table. The ECG fluctuations during the experimental session were visualised in the figure.

Table 8.7: Results of multiple comparisons test with Bonferroni correction for the ECG variables that significantly differed during the time-course of the experimental session.

	Chi^2	P value M1 vs M2	P value M1 vs M3	P value M2 vs M3	M1-baseline mean	M2-after manual mean	M3-after aut. mean	SD for M1	SD for M2	SD for M3	SD for M1, M2, and M3
HR	22.78	<.05	<.05	NS	67.63	65.43	63.75	11.32	10.37	9.77	12.19
HRV low frequencies	15.96	NS	<.05	<.05	2177.3	2195.3	2980.1	4487.44	3112.54	3887.49	4586.47

	Chi^2	P value M1 vs M2	P value M1 vs M3	P value M2 vs M3	M1-baseline mean	M2-after manual mean	M3-after aut. mean	SD for M1	SD for M2	SD for M3	SD for M1, M2, and M3
HRV high frequencies	9.53	<.05	<.05	NS	2903.2	2933.8	3029.8	6631.83	6621.39	5460.99	6520.28
HRV sympathetic tone	7.60	NS	<.05	NS	0.48	0.49	0.55	0.23	0.23	0.21	0.22
HRV vagal tone	7.60	NS	<.05	NS	0.52	0.51	0.45	0.23	0.23	0.21	0.22
HRV RSA	7.90	NS	<.05	NS	6.91	7.02	7.13	1.42	1.39	1.36	1.40
HRV pNN50%	17.55	<.05	<.05	NS	34.33	38.20	39.23	21.59	23.01	22.38	22.71

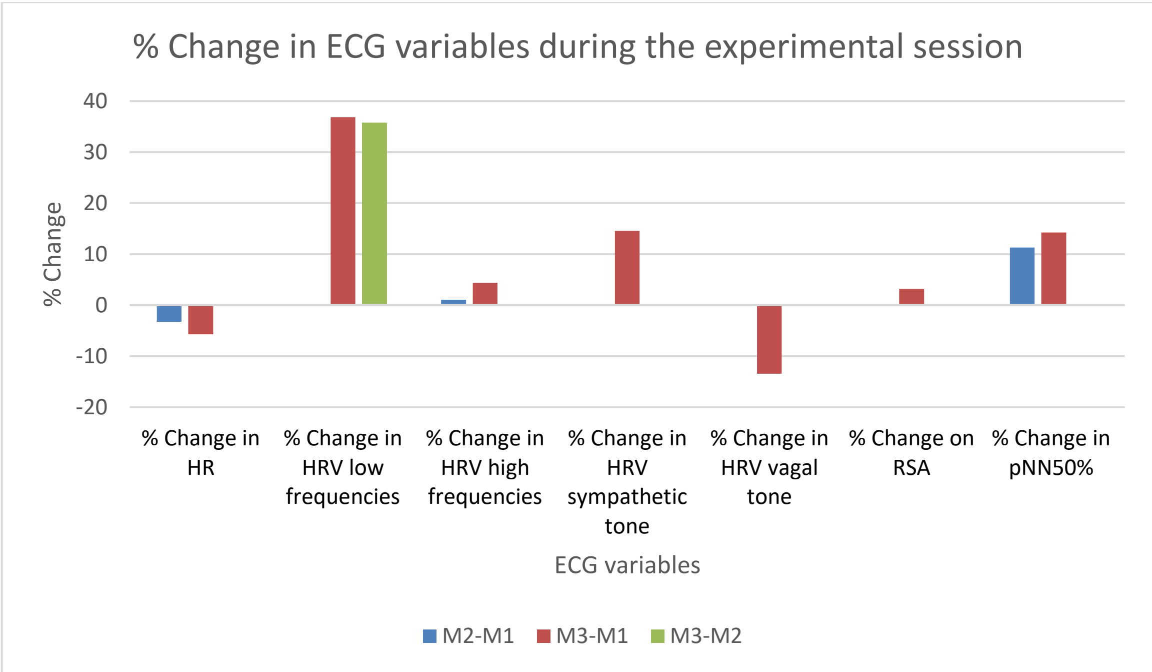

Figure 8.15: Percentage change in the ECG variables that significantly differed between measurement points during the experimental session: blue bars represent change between baseline measurement and measurement after manual driving, red bars change between baseline measurement and measurement after automation, and green bars change between measurement after manual driving and measurement after automation. Positive values represented the percentage increase and negative values percentage decrease in the ECG variables.

8.3.4 VOICE

To investigate the changes in voice over the experimental session, the Kruskal-Wallis test was conducted to compare each voice variables during M1, M2 and M3. The comparison was conducted on the centralised values of the variables to combine both within and between subjects' information. In the case of significant results, multiple comparisons test was conducted for further investigation. The results described below included only voice variables that significantly differed between M1, M2, and M3.

The number of pulses, defined as large maxima in harmonic voice signal (Boyanov & Hadjitodorov, 1997), significantly differed between the measurement points. A multiple comparisons test showed that the number of pulses during baseline measurement was not significantly different from the number of pulses after manual driving, but there was a significant decrease after the automation.

The number of periods, defined as the number of cycles in a given time (Boyanov & Hadjitodorov, 1997), significantly differed between the measurement points. A multiple comparisons test showed that the number of periods during baseline measurement was not significantly different from the number of periods after manual driving, but there was a significant decrease in the number of periods after automation.

The number of breaks, defined as unvoiced segments of the speech (Boyanov & Hadjitodorov, 1997), significantly differed between the measurement points. A multiple comparisons test showed a significant decrease between the baseline measurement and the measurement after automation.

Shimmer, defined as irregularities in voice amplitude (van Lieshout, 2003), significantly differed between the measurement points. There was a significant increase in shimmer between the baseline measurement and the measurement after the automated mode.

The mean autocorrelation, defined as the mean correlation between voice signals separated with a unit of time (Gibbon et al., 1997), significantly differed between the measurement points. A multiple comparisons test showed that mean autocorrelation significantly decreased between the baseline measurement and the later measurements but not between the measurement after manual driving and after automation.

The noise to harmonics ratio, defined as the ratio of noise to harmonics sound energy (Maryn et al., 2010), and related to its harmonics to noise ratio, significantly differed between the measurement points. A multiple comparisons test showed that the mean noise to harmonics ratio significantly decreased between baseline measurement and later measurements but not between the measurement after manual driving and after automation.

The mean intensity, where speech intensity was defined as the squared amplitude of the voice from the beginning of the period until the given point (Gibbon et al., 1997), significantly differed between the measurement points. A multiple comparisons test showed that mean intensity significantly decreased between baseline measurement and later measurements but not between the measurement after manual driving and after automation.

The maximal intensity significantly differed between the measurement points. A multiple comparisons test showed that maximal intensity significantly decreased between baseline measurement and later measurements but not between the measurement after manual driving and after automation.

The detailed results of multiple comparisons tests were presented in table 8.10. The percentage of voice fluctuations during the experimental session was visualised in figure 8.15.

Table 8.8: Results of multiple comparisons test with Bonferroni correction for the voice variables that significantly differed during the time-course of the experimental session.

	Chi^2	P value M1 vs M2	P value M1 vs M3	P value M2 vs M3	M1-baseline mean	M2-after manual mean	M3-after aut. mean	SD for M1	SD for M2	SD for M3	SD for M1, M2, and M3
Number of pulses	23.05	NS	<.05	NS	248.16	242.16	230.43	88.28	90.55	77.36	84.43
Number of periods	21.34	NS	<.05	<.05	231.52	226.77	215.95	87.75	84.23	75.25	81.44
Number of breaks	8.90	NS	<.05	NS	9.99	9.41	9.07	4.72	2.28	2.11	3.05
Shimmer	11.55.	NS	<.05	NS	9.84	10.31	10.60	3.69	3.86	3.95	3.78
Mean autocorrelation	14.27	<.05	<.05	NS	.8554	.8553	.8514	0.07	0.04	0.04	0.05
Mean noise to harmonics ratio	12.16	<.05	<.05	NS	0.29	0.21	0.22	0.86	0.06	0.08	0.46
Mean harmonics to noise ratio	23.41	<.05	<.05	NS	10.98	10.61	10.41	2.20	2.06	2.23	2.13
Mean intensity	18.34	<.05	<.05	NS	61.24	60.25	60.08	5.55	5.97	6.30	5.85
Maximal intensity	16.44	<.05	<.05	NS	75.57	70.42	69.36	41.87	6.08	8.65	22.93

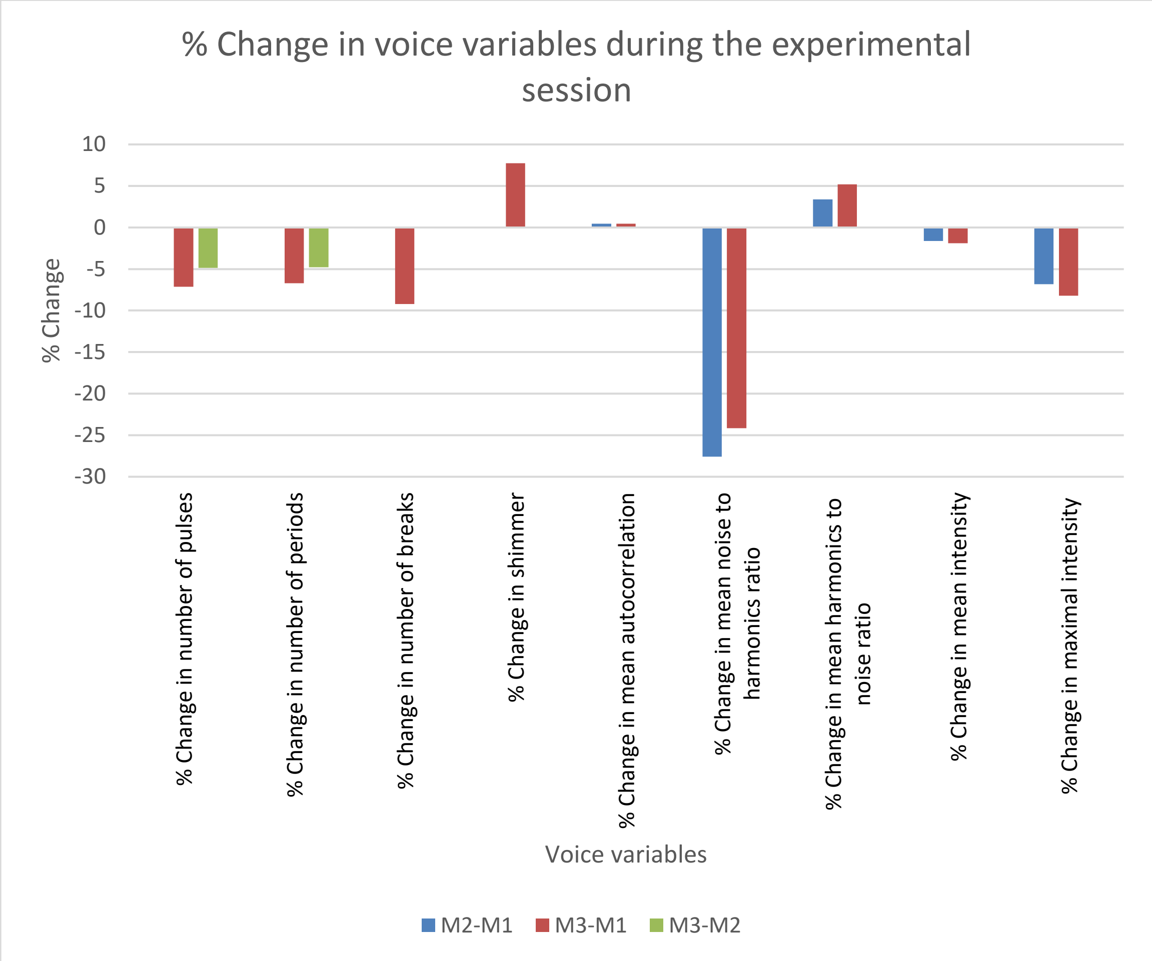

Figure 8.16: Percentage change in the voice variables that significantly differed between measurement points during the experimental session: blue bars represent change between baseline measurement and measurement after manual driving, red bars change between baseline measurement and measurement after automation, and green bars change between measurement after manual driving and measurement after automation. Positive values represented the percentage increase and negative values percentage decrease in the voice variables.

8.3.5 ELECTROOCULOGRAPHY

To investigate the changes in EOG over the experimental session, the Kruskal-Wallis test was conducted to compare each EOG variables during M1, M2 and M3. The comparison was conducted on the centralised values of the variables to combine both within and between subjects' information. In the case of significant results, Bonferroni corrected multiple comparisons test was conducted for further investigation. The results described below include only EOG variables that significantly differed between M1, M2, and M3.

Mean Blink Duration significantly differed between the measurement points ($chi^2(2,46)=7.29$, $p<.05$). A multiple comparisons test showed that mean Blink Duration during baseline measurement ($M = 0.15$, $SD = 0.09$) was not significantly different from the mean Blink Duration after manual driving ($M = 0.17$, $SD = 0.1$), but significantly different than the mean Blink Duration after automation ($M = 0.21$, $SD = 0.14$). The difference in mean Blink Duration after manual driving and automation was non-significant.

8.3.6 ELECTRODERMAL ACTIVITY

To investigate the changes in EDA over the experimental session, the Kruskal-Wallis test was conducted to compare each EDA variables during M1, M2, and M3. The comparison was conducted on the centralised values of the variables to combine both within and between subjects information. In the case of significant results, multiple comparisons test was conducted for further investigation. The results described below include only EDA variables that significantly differed between M1, M2, and M3.

Mean SCL significantly differed between the measurement points (*chi^2(2,85)=73.11*, *p<.05*). A multiple comparisons test showed that mean SCL has significantly decreased between the baseline measurement ($M = 7.39$, $SD = 5.15$) and the measurement after automation ($M = 8.63$, $SD = 5.62$), but not between the baseline measurement and measurement after a manual drive ($M = 9.88$, $SD = 5.62$) or measurement after manual drive and measurement after automation.

8.3.7 ELECTROMYOGRAPHY

To investigate the changes in EMG over the experimental session, the Kruskal-Wallis test was conducted to compare each EMG variables during M1, M2 and M3. The comparison was conducted on the centralised values of the variables to combine both within and between subjects information. In the case of significant results, multiple comparisons test was conducted for further investigation. The results described below included only EMG variables that significantly differed between M1, M2, and M3.

Mean power in corrugator supercilii significantly differed between the measurement points. The mean power significantly decreased between the baseline measurement and the measurement after automation, but not between measurement after manual driving and after automation or at baseline and after manual driving.

Total power in corrugator supercilii significantly differed between the measurement points. The same as mean power it also significantly decreased only between the baseline measurement and the measurement after automation.

The table listed all the details of the multiple comparisons test. The figure presented the percentage of significant changes between the measurement points in the EMG variables.

Table 8.9: Results of multiple comparisons test with Bonferroni correction for the EMG variables that significantly differed during the time-course of the experimental session.

	Chi^2	P value M1 vs M2	P value M1 vs M3	P value M2 vs M3	M1- baseline mean	M2- after manual mean	M3- after aut. mean	SD for M1	SD for M2	SD for M3	SD for M1, M2, and M3
Mean Power Corrugator Supercilii	8.26	NS	<.05	NS	43.12×10^{-12}	131.68×10^{-12}	31.96×10^{-12}	0.0001×10^{-6}	0.0007×10^{-6}	0.0001×10^{-6}	0.0004×10^{-6}
Total Power in Corrugator Supercilii	3.38	NS	<.05	NS	353.29×10^{-9}	1.0788×10^{-6}	261.85×10^{-9}	0.74×10^{-6}	5.73×10^{-6}	0.63×10^{-6}	3.60×10^{-6}

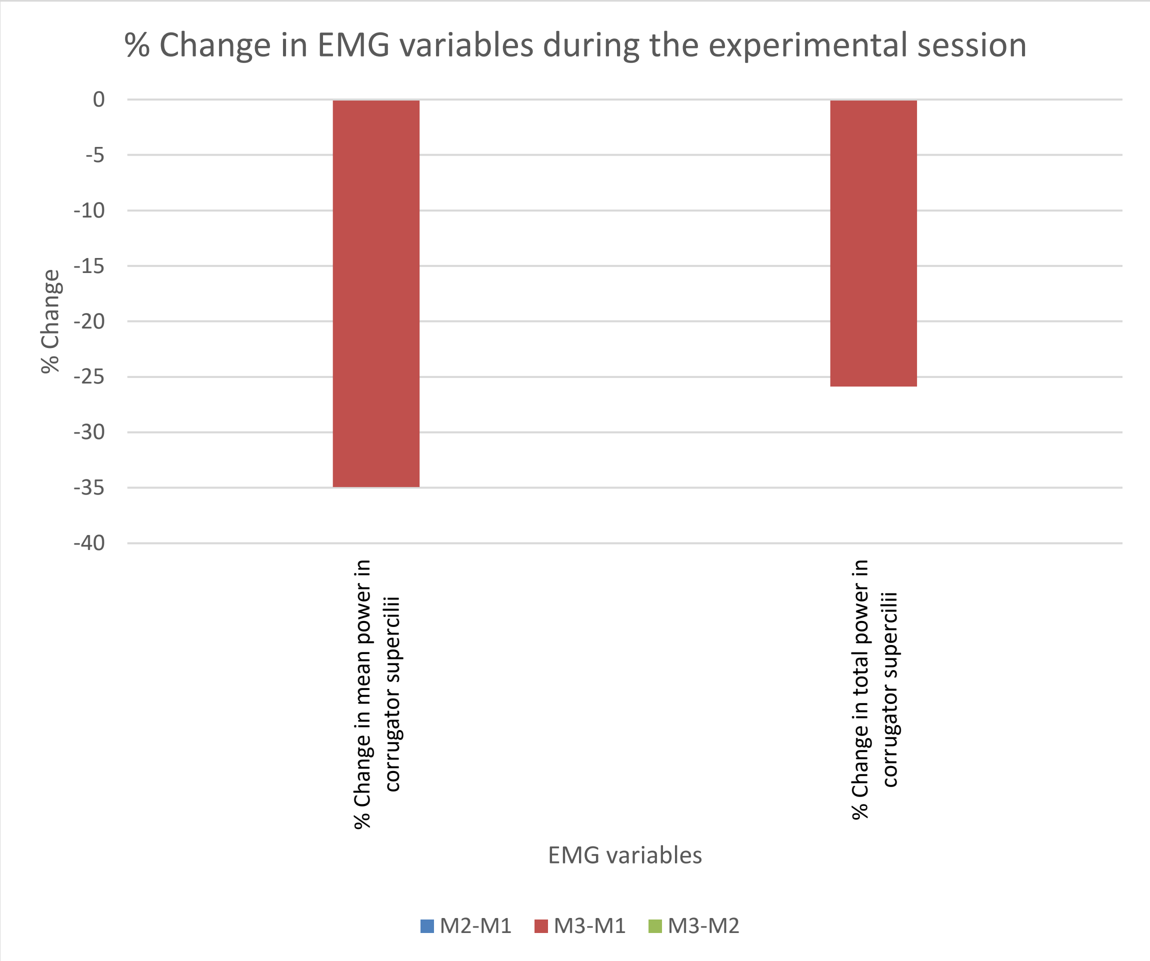

Figure 8.17: Percentage change in the EMG variables that significantly differed between measurement points during the experimental session: blue bars represent change between baseline measurement and measurement after manual driving, red bars change between baseline measurement and measurement after automation, and green bars change between measurement after manual driving and measurement after automation. Positive values represented the percentage increase and negative values percentage decrease in the voice variables.

8.3.8 RESPIRATION

To investigate the changes in Respiration over the experimental session, the Kruskal-Wallis test was conducted to compare each Respiration variables during M1, M2 and M3. The comparison was conducted on the centralised values of the variables to combine both within and between subjects' information. In the case of significant results, multiple comparisons test was conducted for further investigation. The results described below include only Respiration variables that significantly differed between M1, M2, and M3.

The mean standard deviation of the breath amplitude (SD of breath) significantly differed between the measurement points. A multiple comparisons test showed that the mean SD of breath differed significantly decreased between the baseline measurement (*M = 1.27, SD = 1.21*) and the measurement after manual driving (*M = 1.09*, *SD = 1.54*), but not between these measurements and the measurement after automation (*M = .16, SD = 1.33*).

8.3.9 OXIMETRY

To investigate the changes in oximetry based variables over the experimental session, the Kruskal-Wallis test was conducted to compare each variable during M1, M2 and M3. The comparison was conducted on the centralised values of the variables to combine both within and between subjects' information. In the case of significant results, multiple comparisons test was conducted for further investigation. The results described below included only oximetry variables that significantly differed between M1, M2, and M3.

Mean blood oxygenation significantly differed between the measurement points. A multiple comparisons test showed that mean blood oxygenation increased significantly between the baseline measurement and measurement after automation.

The mean pulse also significantly differed between the measurement points. A multiple comparisons test showed that the mean pulse was significantly decreasing during the experimental session.

Table 8.12 listed all the details of the multiple comparisons test. The figure presented the percentage of significant changes between the measurement points in the oximetry-based variables.

Table 8.10: Results of multiple comparisons test with Bonferroni correction for the EMG variables that significantly differed during the time-course of the experimental session.

	Chi^2	P value M1 vs M2	P value M1 vs M3	P value M2 vs M3	M1-baseline mean	M2- after manual mean	M3-after aut. mean	SD for M1	SD for M2	SD for M3	SD for M1, M2, and M3
Blood Oxygenation	10.16	NS	<.05	NS	97.83%	98.08%	98.09%	1.30	1.47	1.53	2.27
Pulse	35.68	<.05	<.05	<.05	66.96	64.90	64.47	10.19	11.00	16.04	11.56

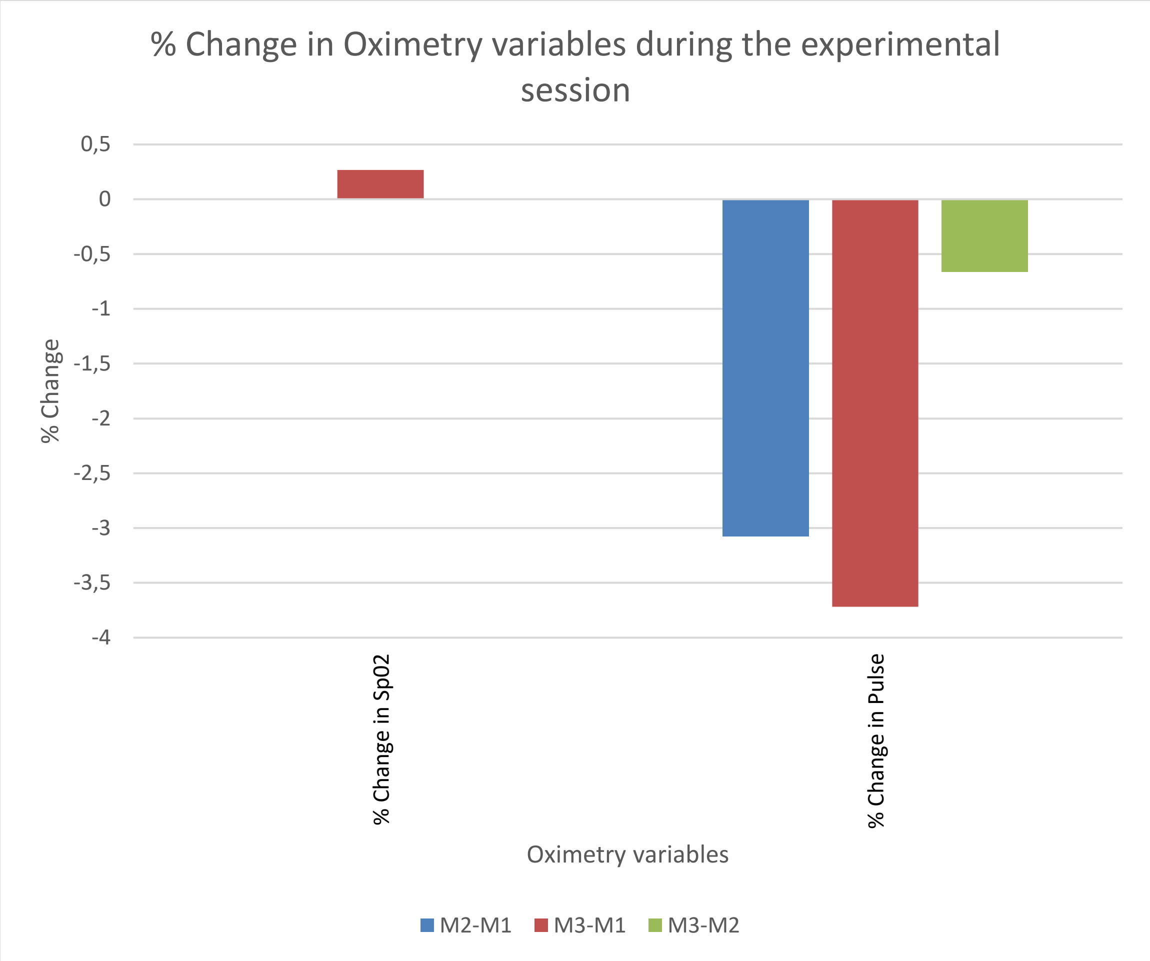

Figure 8.18: Percentage change in the oximetry based variables that significantly differed between measurement points during the experimental session: blue bars represent change between baseline measurement and measurement after manual driving, red bars change between baseline measurement and measurement after automation, and green bars change between measurement after manual driving and measurement after automation. Positive values represented the percentage increase and negative values percentage decrease in the voice variables.

8.4 DISCUSSION

The analysis reported in this chapter compared manual driving performance before and after automation, as well as physiology and subjective psychological state of the participants at the baseline level, after manual driving, and after automation. The hypothesis was supported, suggesting that the driving performance, subjective sleepiness, subjective fatigue, and subjective readiness to take-over manual driving were worse after the driver experienced automation. Also, changes in ECG could indicate lower mental workload after automation, maybe to the extent of the underload. However, it is essential to note that the experiment did not allow dissociation between the effect of automation and the effect of prolonged participation in the experiment. To fully confirm that the effects were related to the automation, not to the time-on-task, another experimental work needs to be conducted with the control group of the same task duration but without the automated mode. Such a control group was not included in this experimental design due to the temporal and financial limitations. It would either require recruiting twice more participants or assign half of the participants to the group without the automation, which, as a result, would reduce the statistical power of the models of take-over and automation monitoring. Experimental work took ten months, and the number of participants was not enough to divide it in half. Due to these reasons, such validation was left for future experimental work.

The experiment might provide some data consistent with previous predictions related to semi-automated driving. Several researchers have suggested that automation might reduce the performance of drivers. The reasons might be the demand for sustained attention, being out of the control group (or in other words not being fully aware of the situation), as well as fatigue (Hancock, 2015; Kyriakidis et al., 2019; Young & Stanton, 2007). Previous results have shown an increase in drowsiness after both manual and automated driving (Schömig et al., 2015). The analysis of the experimental data showed a decrease in driving performance after participants went through the automated phase; however, the current experiment did not provide enough evidence to fully differentiate it from the effects of time on task, especially as manual driving might also increase sleepiness (Schömig et al., 2015).

Moreover, the decrease was more profound during night-time driving than during daytime driving. The negative influence of the night circadian phase confirmed predictions that were previously only formed based on manual driving or various literature reviews. It was previously predicted that changes in attention, fatigue, and other cognitive functions that people experience at night might worsen driver performance in semi-automated driving (Kaduk et al., 2020); however, again it is necessary to differentiate it in future research from the effects of time on task. In the current experiment, driving performance decreased after automation both during day and night experimental sessions; however, more driving factors deteriorated at night. Subsequently, with the decrease in the driving performance, participants also experienced some changes in sleepiness, fatigue, subjective readiness to take-over manual driving, and mental workload after the automated phase. They felt sleepier, more fatigued, and less ready to take-over manual control after the automation. It was also notable that they assessed the mental workload during automation as lower than the mental workload during manual driving. If it were assumed that such a decrease resulted in a mental underload, it would be consistent with the data showing a performance decrease when the mental workload is too low as well as the predictions that automation might decrease the mental workload of the driver (Heikoop et al., 2016; Young & Stanton, 2002, 2007). The increase in sleepiness, fatigue, and feeling of not being ready to drive could also be related to the depletion of mental resources caused by prolonged and sustained attention (Hancock, 2015; Warm et al., 2008).

Interestingly, the subjective state of the participants was not correlated with the changes in driving performance. It indicated that participants' manual driving was poorer after automation and subjectively they felt more tired and less ready to drive. Still, the magnitude and occurrence of these two processes did not significantly correlate. It could be related to several functions, such as subjective sleepiness and fatigue, which are different from biological sleepiness and fatigue. However, the information that might have even more profound consequences for driving was that participants were not able to predict how ready they were to take-over manual driving with the TORS questionnaire. It could indicate that drivers were not able to accurately assess their state or the fitness to drive. Such a lack of awareness could lead to take-overs in inappropriate situations and continuation of driving when drivers should take a break. These findings suggested that there is a requirement for additional methods of driver state monitoring, for example, in-car physiological sensors.

There is an ongoing discussion in the scientific literature about whether driving performance reduces as a result of automation. Many researchers have predicted that automated vehicles could negatively affect manual driving; however, some did not observe any performance decrease (Merat et al., 2014). In the current study, a reduction in driving performance was observed after the automated phase. It is possible that it was an effect of the automated mode; however, it could also be caused by other factors, such as time on task. The reduction in manual driving performance was not restricted to the period just after take-over, as manual driving in the current experiment took place approximately ten minutes after the end of the automated mode. Participants were not accurate in the assessment of their fitness to drive, which made such a risk more profound.

In respect to physiology, there were many changes observed during the time-course of the experimental session. From the perspective of the research questions, the most interesting changes were from the measurements taken after manual and automated driving. The period in between these measurements was mainly filled with a simulation of the automated mode, and hence it was possible that the changes were related to the automation.

Power in low frequencies in heart rate variability significantly increased after automation. This range of frequencies, especially the 0.1 Hz component, were previously associated with the mental effort, mental workload and time on task (Fairclough & Houston, 2004).

The increase in power could indicate decreased mental workload or underload related to automation.

The number of pulses and the number of periods significantly decreased between the measurement after manual driving and measurement after automation. They were not previously used in state monitoring.

Pulse has significantly decreased between the measurement after manual driving and after automation. Sanpeng et al. (2010) reported that increased pulse could be an indicator of extreme fatigue; however, the pulse is a highly unspecific method and it is hard to draw precise psychological or behavioural conclusions based on it.

This study was not without the limitations that should be understood as a part of data interpretation. The experimental design did not allow dissociation between the effect of automation and time on task. The sample size in the circadian analysis was too small to draw definite conclusions. Also, it should be noted that the standard deviation of some of the driving performance measures was relatively high and could lead to confounding effects, despite the statistical significance of the results. There is a need for future work to control for these variables.

8.5 CONCLUSIONS

This study indicated certain risks for manual driving after the preceding automated phase. People generally seemed to drive worse after using vehicle automation. They also reported feeling more fatigued, sleepy, and less ready to drive after using automation, particularly at night. Their heart rate variability also suggested mental underload related to vehicle automation. As semi-automated vehicles are entering the marketplace, it is, therefore, crucial to understand the mechanisms that cause driver performance decrements. As people are not good at assessing their fitness to drive, manufacturers cannot solely rely on the judgment of the driver. Potential mitigating measures could include psychophysiological monitoring systems in the vehicle that would assess the driver state as well as a change in the design of automation to better interact with human cognition.

9. CIRCADIAN EFFECT ON SEMI-AUTOMATED DRIVING AND DRIVER STATE MONITORING- EXPERIMENTAL RESULTS

This chapter included the results of the circadian analysis of the experimental data. Literature reviews on circadian effect on manual driving, semi-automated driving, and driver state monitoring suggested an adverse effect of night-time on driving performance and changes in the interpretation of physiological data based on the circadian phase. The following subchapters described experimental validation of these theoretical predictions.

9.1 OVERVIEW OF THE GENERAL EXPERIMENTAL METHODS AND THE STUDY SPECIFIC DIFFERENCES

Chapters 4, 5, and 6 showed evidence that the circadian phase might affect driving performance in manual and semi-automated driving, as well as should be included in the systems of driver state monitoring to assure better accuracy and interpretation of the signals. The following sections described experimental work and analysis of the circadian effect on simulated semi-automated driving. The main hypotheses for this chapter were that there is a circadian effect on semi-automated driving and psychophysiology of the driver in the semi-automated vehicle. The sub-hypotheses were that driving performance and attention during automation decrease during the night.

The experimental design, participants and set-up were already described in chapter 8.2; however, this section presented the perspective of the circadian effect.

Data collected from fifty-two participants were analysed in the circadian context. The analysed variables were related to driving performance, attention during automation, EMG, EOG, ECG, respiration, EDA, voice, oximetry, salivary cortisol and alpha-amylase, subjective sleepiness measured with KSS, subjective readiness to take-over measured with TORS, subjective fatigue measured with Samn-Perelli Fatigue scale, and subjective mental workload measured with NASA-TLX.

The driving tasks were performed in the low-fidelity driving simulator with the STISIM 3 software. The simulator was placed inside the noise-insulated Faraday cage,

The analysed psychophysiological variables came from the periods of resting-state during M1, M2, and M3, the manual driving performance came from the driving tasks T1 and T2, and the attention measure came from the period of automated mode, as shown in figure 9.1. The same experimental procedure was repeated twice for each participant, once during the high day-time circadian phase (9 a.m.- 1 p.m.) and once during the low night-time circadian phase (10 p.m.- 2 a.m.).

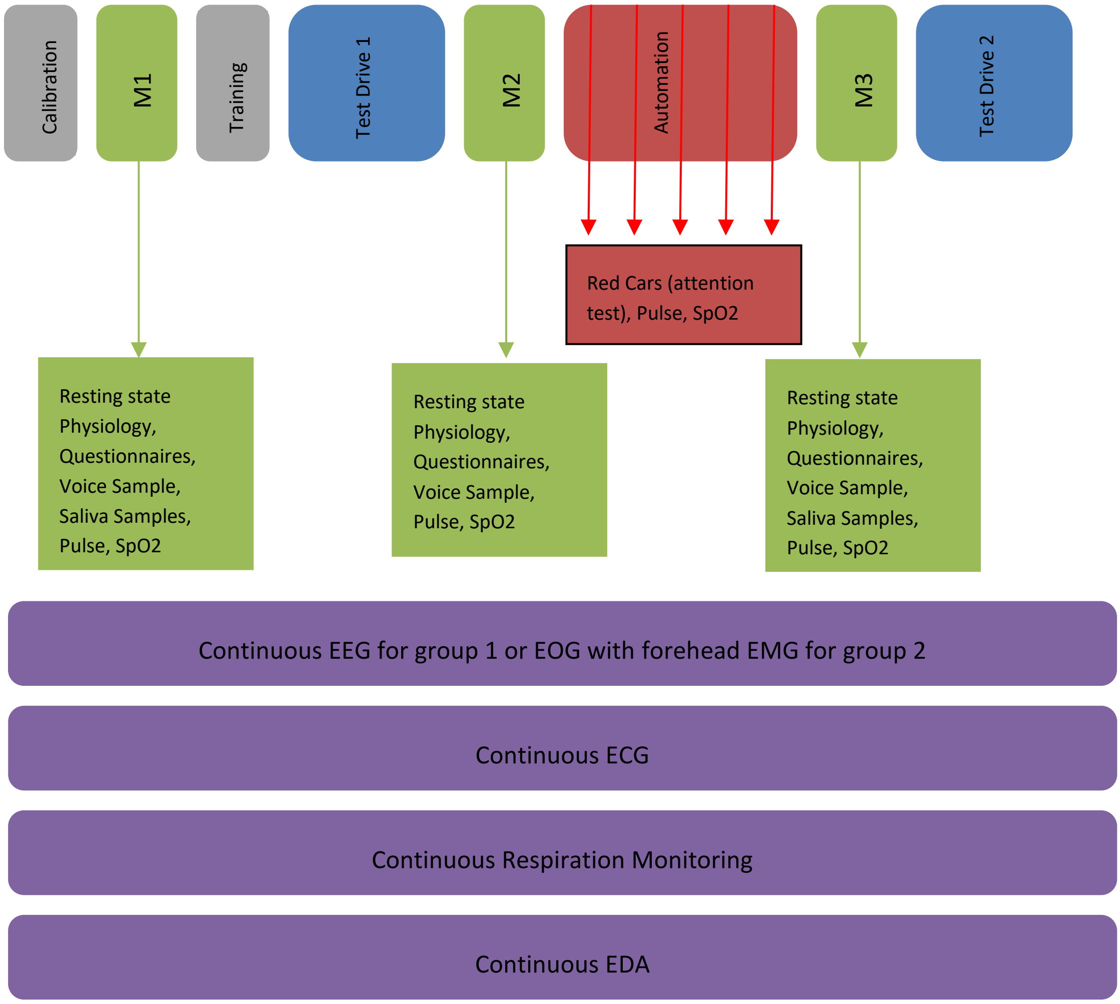

Figure 9.1: The outline of the experimental design.

9.1.1 STATISTICAL DATA ANALYSIS

The data analysis consisted of three major steps: data pre-processing, signal analysis and statistical analysis. As pre-processing and signal analysis did not differ for the circadian analysis, only the statistical analysis was described in this chapter.

The Kolmogorov-Smirnov test showed that none of the experimental variables had a normal distribution, hence only non-parametric correlation tests were used.

To compare different factors between the circadian phases, the rho-spearman correlation was calculated between the circadian phase (binary variable with values 0 for the day and 1 for the night) and different psychophysiological variables. The analysis of the psychophysiological variable was conducted for the resting state measurements M1, M2, and M3. Only the factors that significantly correlated with the circadian phase were described in the next subchapter. The statistical analysis and data processing were conducted in Matlab R2020a and Excel.

9.3 RESULTS OF THE EXPERIMENT

9.3.1 MANUAL DRIVING PERFORMANCE

The Bonferroni corrected rho-Spearman correlation was calculated between driving performance and circadian phase. None of the correlations reached the significance level.

The first experimental session was characterised by a stronger learning effect on the driving performance as participants were doing the driving tests for the first time, and most of them were using the driving simulator for the first time. Because of that, analysis of the association between circadian phase and driving performance was conducted again separately for the first and second session.

During the first experimental session, there was a significant positive correlation between the number of collisions and circadian phase ($r_s(100) = .29$, $p<.05$). While for the second experimental session, the association between the number of collisions and the circadian phase was also significant but with an opposite direction ($r_s(93) = -.21$, $p<.05$). The association was visualized in figure

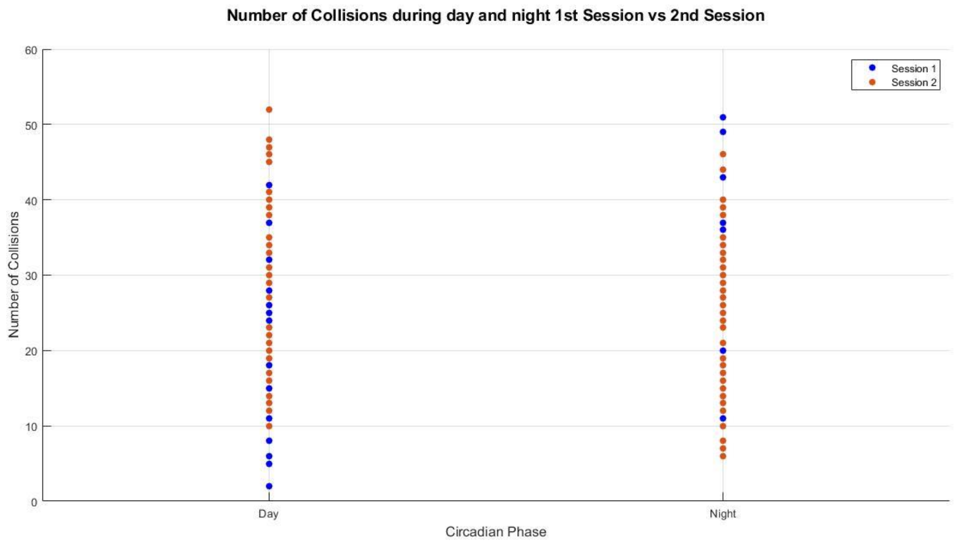

Figure 9.2: Association between the number of collisions and circadian phase during the first and the second experimental session.

As the previous research in the forced desynchrony protocol showed that circadian phase affected driving performance only in interaction with the sleep deprivation (Matthews, Ferguson, Zhou, Kosmadopoulos, et al., 2012; Matthews, Ferguson, Zhou, Sargent, et al., 2012) a stepwise regression was tested using the mean of general driving performance as a dependent variable and circadian phase, the number of hours slept before the experiment, and mean of KSS results as independent variables. P value and adjusted r-squared were used as criteria for the addition or removal of the factors. All the independent factors were removed and the model was insignificant (*F(1,97) = 1.24, adjusted r-squared =NA, p:NS*).

9.3.2 QUESTIONNAIRES

Participants reported their subjective sleepiness, fatigue, readiness to take-over manual driving and mental workload using questionnaires. The questionnaires' results were correlated with circadian phase. As shown in figures 9.2, 9.3, 9.4, and 9.5 scores that significantly correlated with circadian phase were KSS ($rs(301) = .33, p<.05$), TORS ($rs(271) = .24, p<.05$), Fatigue Questionnaire ($rs(296) = .40, p<.05$) and NASA-TLX physical demand scale ($rs(265) = .06, p<.05$).

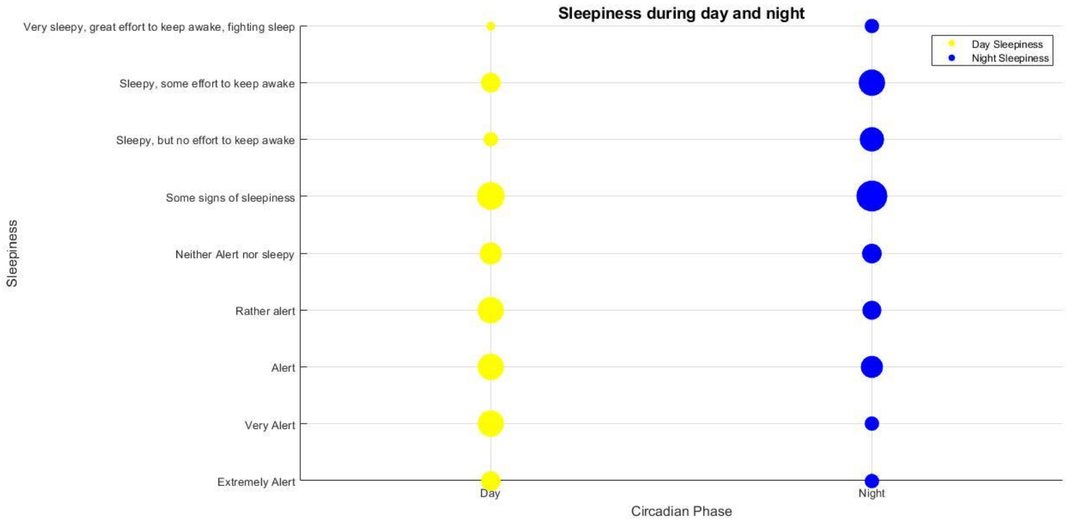

Figure 9.3: Sleepiness increased at night. The size of the markers represents the frequency of the values.

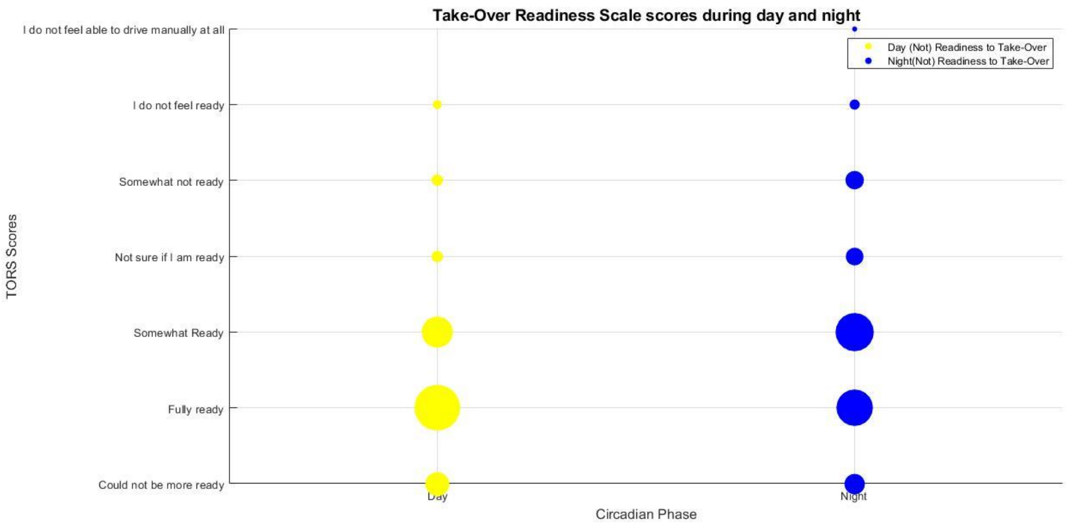

Figure 9.4: Participants felt less ready to take-over at night. The size of the markers represents the frequency of the values.

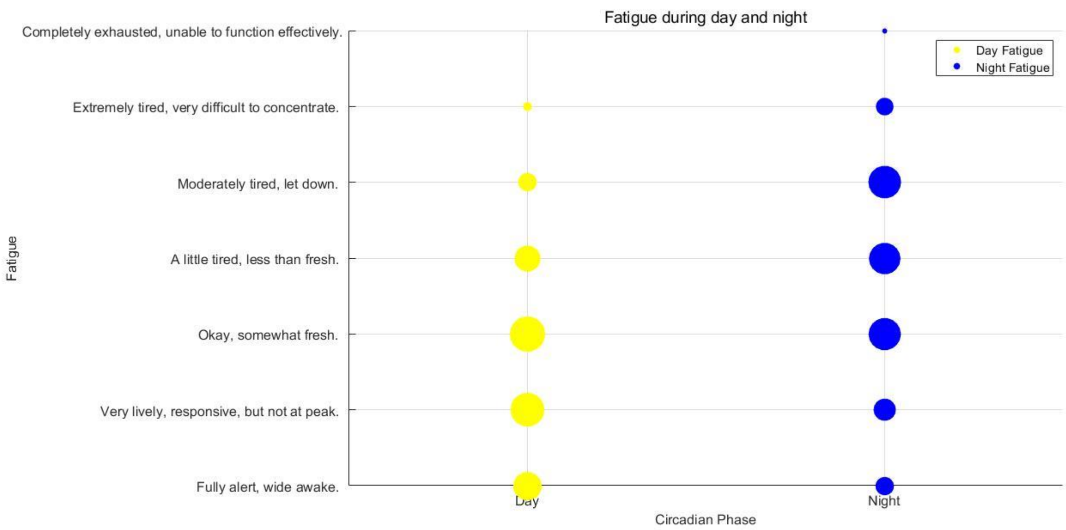

Figure 9.5: Fatigue increased at night. The size of the markers represents the frequency of the values.

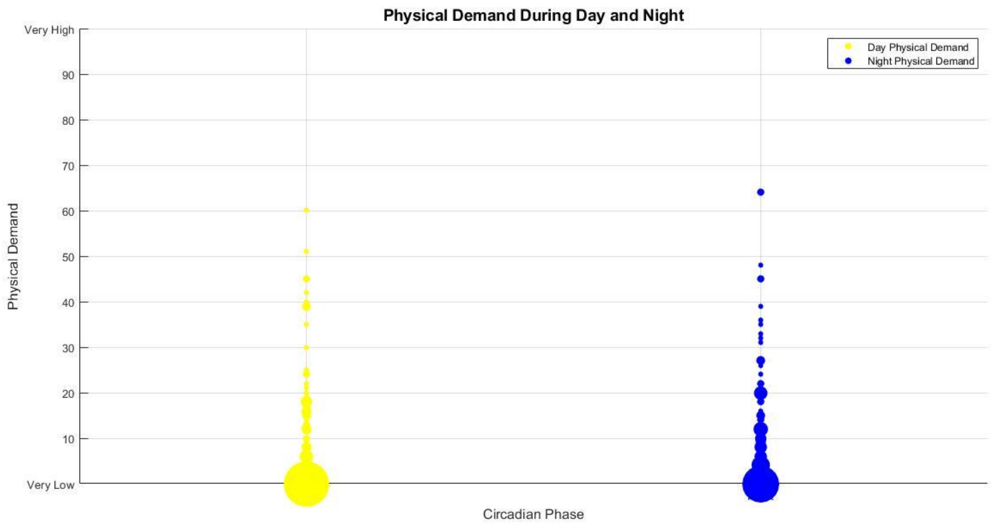

Figure 9.6: Physical demand had a very low tendency to increase at night. The size of the markers represents the frequency of the values.

9.3.3 VOICE

Rho-Spearman correlation test between the Voice measures and a circadian phase (0=day, 1=night) showed a small significant correlation between circadian phase and voice mean autocorrelation (*rs(299) = .2, p<.05*), and between circadian phase and mean noise to harmonics ratio (*rs (299) = -.2, p<.05*) as shown in figure 9.7 and 9.8

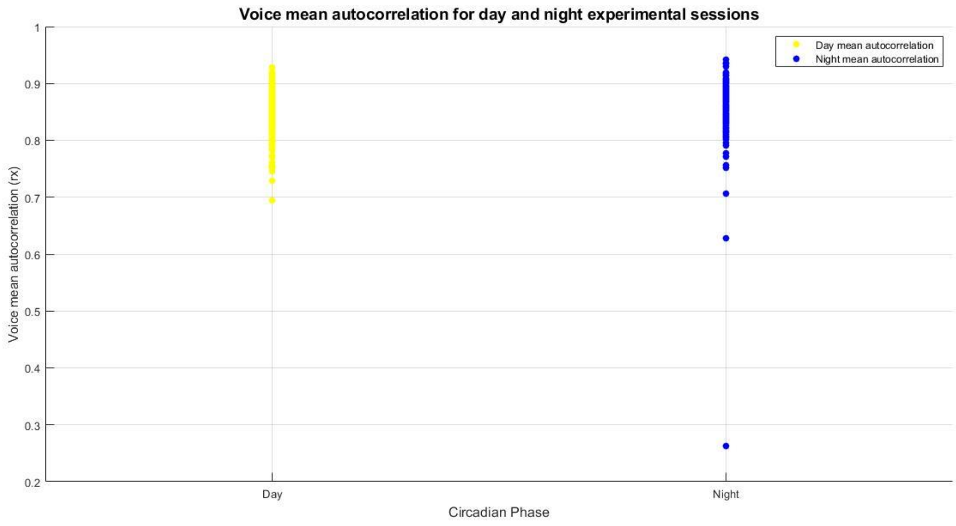

Figure 9.7: Mean voice autocorrelation for day and night experimental sessions. The formula to calculate Rx was provided in a glossary of terms (Yoon et al., 2006).

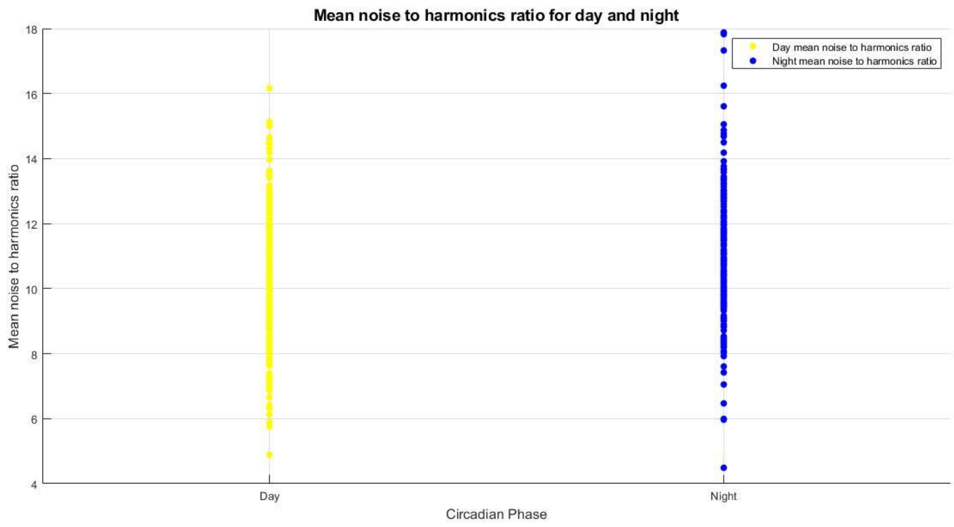

Figure 9.8: Mean voice noise to harmonics ratio for day and night experiments.

9.3.4 ELECTROOCULOGRAPHY

The rho-Spearman correlation was used to test the association between circadian phase and EOG variables. There was a significant correlation with the rate of horizontal eye-movements (*rs (145) = -.22, p<.05*), as shown in figure 9.9.

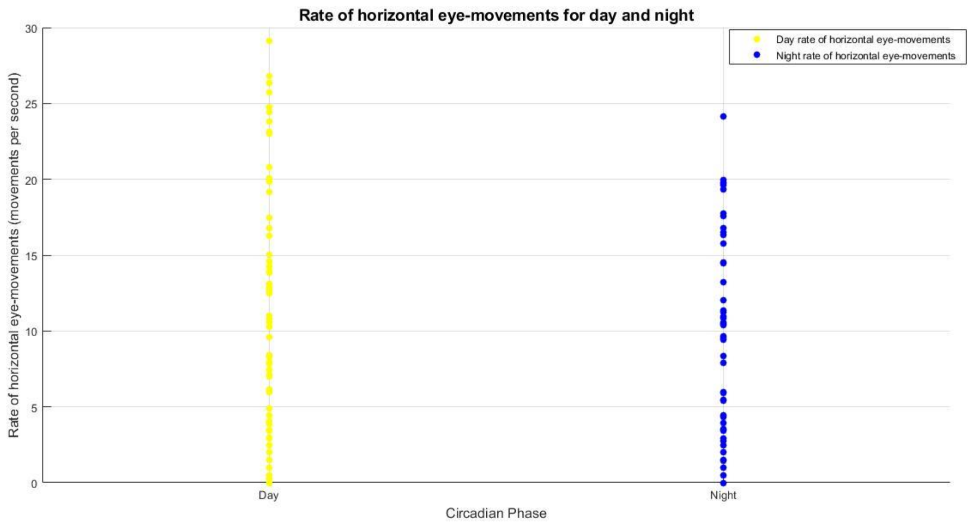

Figure 9.9: Rate of horizontal eye-movement decreased at night.

9.3.5 ELECTROMYOGRAPHY

The rho-Spearman correlation was used to test the association between circadian phase and EOG variables. There was a significant correlation with the mean frequency in frontalis muscle (*rs (147) = -.18, p<.05*), mean power in frontalis muscle (*rs (147) = .18, p<.05*), and total power in frontalis muscle (*rs (147) = .18, p<.05*) as shown in figure 9.10, 9.11 and 9.12.

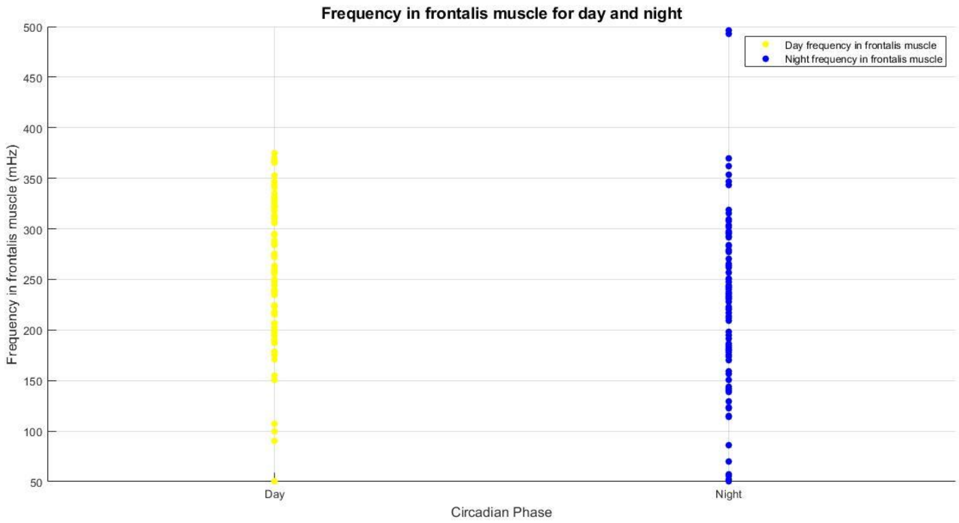

Figure 9.10: Frequency in frontalis muscle tended to decrease at night.

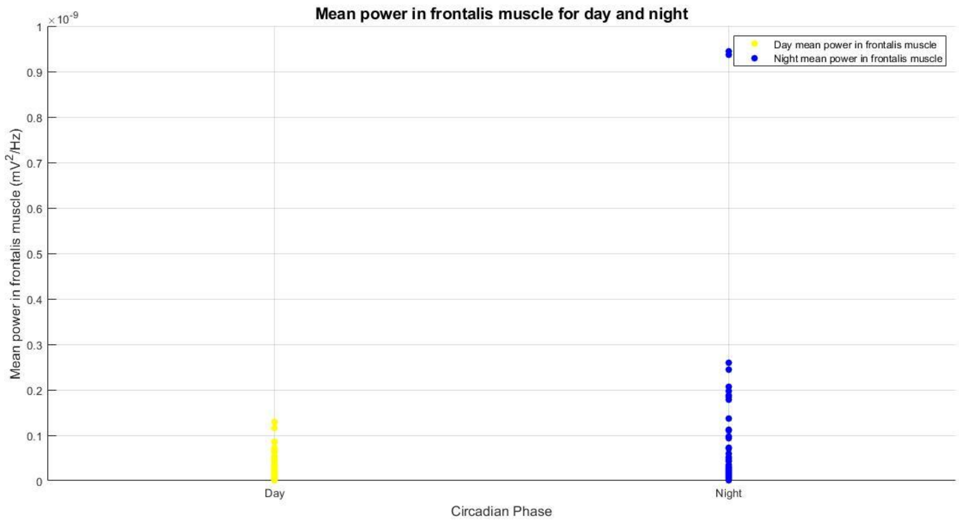

Figure 9.11: Mean power in frontalis muscle tended to increase at night.

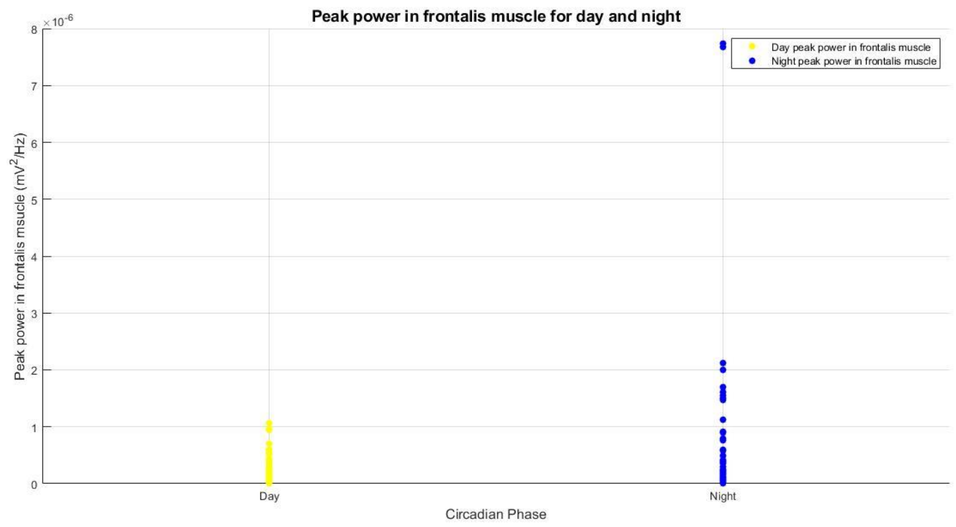

Figure 9.12: Peak power in frontalis muscle tended to increase at night.

9.3.6 SALIVA

Rho-Spearman correlation between circadian phase and cortisol salivary content was significant and high negative (*rs(206)=-0.76, p<.05*) as shown in figure 9.13. In contrast, the correlation with the alpha-amylase content was insignificant.

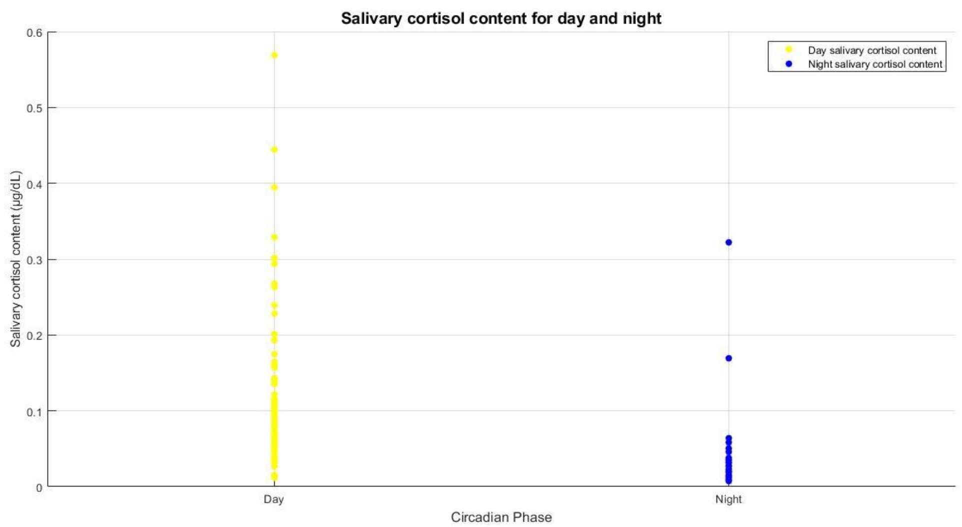

Figure 9.13: Cortisol level was higher during the day.

9.4 DISCUSSION

This chapter investigated the circadian effect on semi-automated driving and driver state monitoring. The literature reviews presented in chapter 4, 5, and 6 provided data about the potential circadian effect on semi-automated driving and driver state monitoring. It allowed formulating of hypotheses for this experiment. The main hypotheses for this chapter were that there is a circadian effect on psychophysiology and performance of the driver in the semi-automated vehicle. The sub-hypotheses was that driving performance and attention during automation decrease during the night. The main hypothesis was supported, but the sub-hypothesis failed to match experimental evidence.

To investigate the circadian effect on psychophysiology and semi-automated driving, the binary variable representing the circadian phase (0-day, 1-night) was correlated with driving performance, attention during automation, and psychophysiological variables. Driving performance, questionnaires' results, voice, EOG, EMG, and salivary cortisol content were found to change with the circadian phase; however, only cortisol level displayed a high correlation. These changes supported the main hypotheses that there are changes in driving performance and psychophysiology of the driver in a semi-automated vehicle.

Driving performance did not differ significantly between day and night session; however, when divided into the first and second experimental session, the number of collisions correlated on the low level with a circadian phase. During the first experimental session, which was characterised by a lower familiarity with the tasks, there were more collisions at night. However, during the second experimental session, when the tasks were more familiar, there were fewer collisions at night. Such results did not support the sub-hypotheses and seemed to contradict the previous literature (Akerstedt et al., 2001; Chipman & Jin, 2009; Mitler et al., 1988). The decrease of the driving performance at night during the first experimental session could be explained by the higher cognitive demand of the unfamiliar tasks; however, it requires further investigation to explain the increase of the driving performance at night during the second session. Future research with different types of driving tasks could give a better understanding of this finding.

The subjective state of the participants differed between the day and night. They declared being sleepier, more fatigued and less ready to take-over manual driving at night than during the day. There was also a small increase in the subjective physical demand of the tasks at night, but the correlation was very low. Higher sleepiness and higher fatigue at night were already researched in the past (Lowden et al., 2009; Otmani et al., 2005); however, lower readiness to take-over is a newly reported phenomenon. It is a factor that could negatively influence the comfort of the night-driving in semi-automated vehicles and lead to disuse or misuse of automation at night.

Two acoustic properties of voice displayed changes during the circadian cycle. Mean autocorrelation showed a small tendency to increase at night, while mean harmonics to noise ration to decrease at night. The tendency was small but statistically significant.

Neither of these properties was reported to be used in state monitoring; however, both increase of autocorrelation and a decrease of harmonics to noise ratio were associated with negative emotions or negative arousal (Linhart et al., 2015; Mongia & Sharma, 2014).

Mean frequency in frontalis muscle tended to increase at night, while mean power and peak power in frontalis muscle tended to decrease at night. This result might have consequences for some systems of driver state monitoring as the increased power in frontalis muscle was previously reported to be associated with increased mental workload (Cohen et al., 1992). Considering that there were almost no circadian differences in mental workload, the changes in frontalis muscle could be addressed to the circadian phase. In the case of driver state monitoring with the use of EMG of the frontalis muscle, this effect should be taken into account.

There was a tendency to the decreased rate of the horizontal eye-movements at night. Previous literature identified a decrease in spontaneous and saccadic eye-movements as an indicator of increased mental workload (May et al., 1990) as well as with increased fatigue (Lal & Craig, 2001, 2002) and drowsiness (Borghini et al., 2014). It would require additional experimental work to dissociate the effect of the circadian phase from the influence of drowsiness and fatigue on horizontal eye-movements as both fatigue and drowsiness increased at night. If there was a circadian change in the rate of horizontal eye-movements, it would be recommended to treat the circadian phase as a factor in driver state monitoring systems that use ocular behaviours.

Cortisol level in saliva significantly decreased at night, which is a result repeatedly reported before. Cortisol is treated as one of the main hormonal components of the circadian cycle, together with melatonin (Blatter & Cajochen, 2007; Cajochen et al., 2002).

These results allowed to support the suggestion presented in the review in chapter 6, that the circadian phase might influence systems of driver state monitoring. However, it was not confirmed that the driving performance in semi-automated vehicle decreased at night, which was suggested in the theoretical chapter 5.

9.5 CONCLUSIONS

The circadian phase can influence the driver psychological and physiological state. Such influences should be taken into account when designing a semi-automated system.

In the design process of the systems of driver state monitoring, the circadian effect should be tested as one of the variables, particularly as some of the factors in driving performance might get worse at night.

Increased feeling of sleepiness, fatigue, and decreased readiness to take-over manual driving at night might reduce the willingness of drivers to engage in a semi-automated driving system. The design of automated systems should anticipate the potential states of the driver at night.

To conclude, the circadian effect tends not to be included in the design of semi-automated vehicles, but it is worth considering the potential effect on a driver's state and physiology. Clearly then, more research is urgently needed.

10. DRIVER STATE MONITORING IN SEMI-AUTOMATED VEHICLES- EXPERIMENTAL RESULTS

The two literature reviews described in chapters 2 and 3, allowed to identify a list of risky driver states and a variety of psychophysiological measures that could detect them. The main question of this doctoral project was what are the most accurate methods to monitor driver state to ensure safe take-over. It would mean monitoring of the driver attention during automated mode to ensure that they are monitoring the process, as well as the short measurement of the driver state before the take-over to check if they are ready to safely resume manual control over the vehicle. To answer these questions, the experiment was conducted in the driving simulator using a high amount of psychophysiological measures. The choice of the measures and the laboratory environment were described in Chapter 7. The experiment was conducted in the laboratory to ensure the precision of the physiological measurements (see chapter 7); however, there was also a need for a certain ecological validity. Many studies investigated driver state monitoring comparing physiological indicators to the questionnaires or experimentally inducing sleepiness and observing accompanying physiological changes (Jiao & Lu, 2016; Samn & Perelli, 1982). This experiment examined the potential to predict actual driving performance or performance in attention task with physiological measures. It was a novel step further into the application of the current knowledge about physiology. However, it also created a more challenging situation, as driving performance does not always change with the state of the driver. For example, a sleepy or distracted driver might still drive well using their additional available cognitive capacity (Parasuraman et al., 2008; Ross et al., 2014; Young & Stanton, 2002). Another challenge was that physiological states often are not unique for some mental states. For example, increased heart rate might be related to any type of increased arousal not being selective for one specific mood. At the same time, it might be associated with the change of bodily position, hence indicate more physical arousal. The following chapters described the methods and results of the experiment. The main hypotheses of this analysis were that 1. psychophysiological measures can predict performance after take-over and 2. Psychophysiological measures can predict attention during automation. The sub-hypotheses was that psychophysiological indicators that can predict driving performance and attention are related to one of the risky states identified in chapter 2, namely sleepiness, fatigue, distraction, mental workload or situation awareness.

10.1 OVERVIEW OF THE GENERAL EXPERIMENTAL METHODS AND THE STUDY SPECIFIC DIFFERENCES

Data collected from fifty-two participants was used to created predictive models of driver performance in the semi-automated vehicle. The independent variables were related to EMG, EOG, ECG, respiration, EDA, voice, oximetry, salivary cortisol and alpha-amylase, subjective sleepiness measured with KSS, subjective readiness to take-over measured with TORS, subjective fatigue measured with Samn-Perelli Fatigue scale, and subjective mental workload measured with NASA-TLX. The dependent variables were related to driving performance and attention during automation.

The driving tasks were performed in the low-fidelity driving simulator with the STISIM 3 software. The simulator was placed inside the noise-insulated Faraday cage,

The analysed psychophysiological variables came from the periods of resting-state during M1, M2, and M3, the manual driving performance came from the driving tasks T1 and T2, and the attention measure came from the period of automated mode, as shown in figure 10.1. The same experimental procedure was repeated twice for each participant, once during the high day-time circadian phase (9 a.m.- 1 p.m.) and once during the low night-time circadian phase (10 p.m.- 2 a.m.).

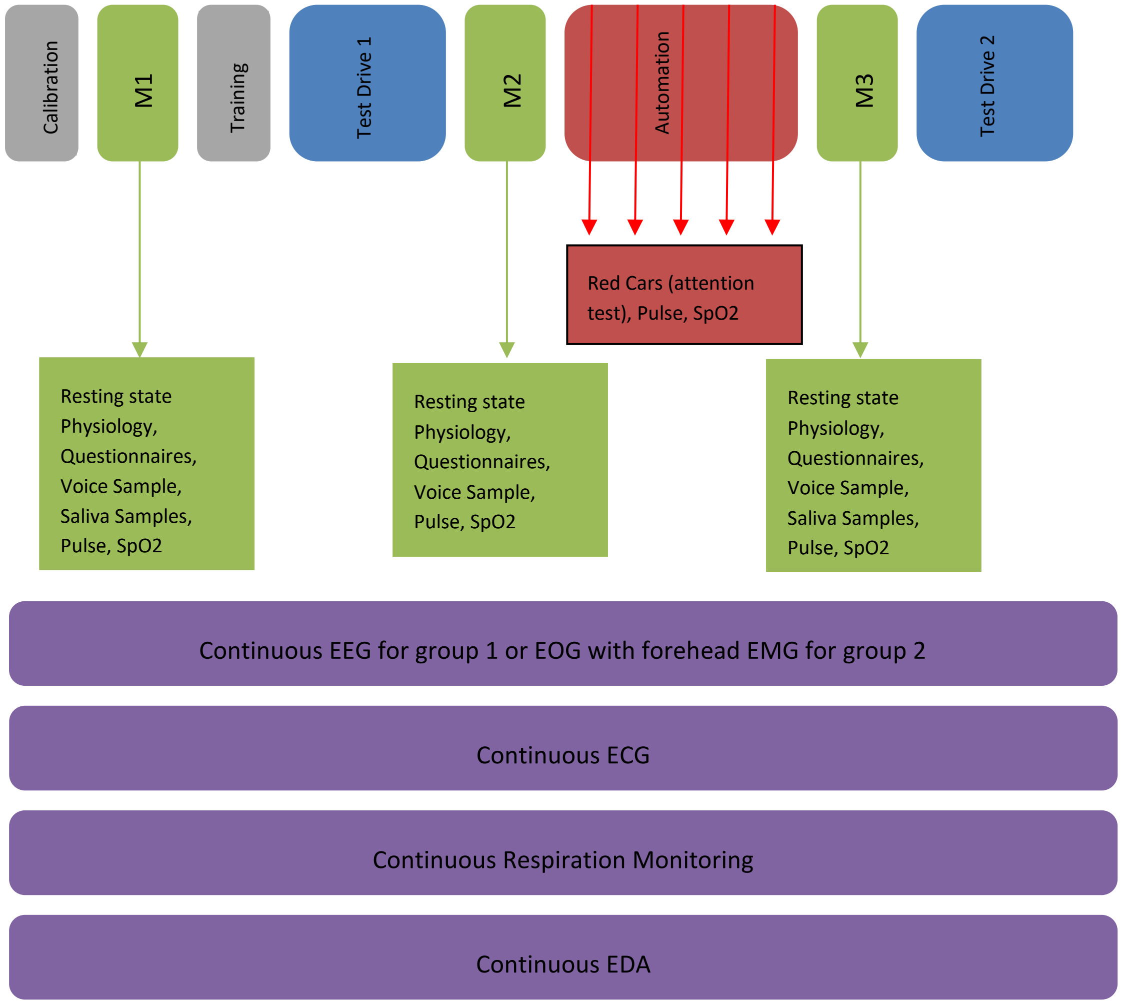

Figure 10.1: The outline of the experimental design.

10.1.1 STATISTICAL ANALYSIS

Each of the physiological variables was tested for normality with the Kolmogorov-Smirnov test. All tests indicated non-normal distributions of the variables, so the statistical tests used in the analyses were non-parametric ones.

To analyse driver state monitoring before take-over new variables were created by subtraction of physiological variables at Measurement 1 and Measurements 3 (M1-M3) and driving performance at Test 1 and Test 2 (T1-T2). It allowed to assess change in physiology and driving performance and avoid confounding results with differences in baseline physiology or baseline driving performance in different subjects. Following this data manipulation linear models were tested using various M1-M3 factors, and some other variables like circadian phase as independent variables and T1-T2 as the dependent variable. Also, the rho-Spearman correlation was calculated between all the physiological variables and different factors of the driving performance. The first step of the analysis was the features selection based on the p values and adjusted r-squared of the linear models using only factors derived from one physiological function as independent variables, as well as rho-Spearman correlation values of the correlation between physiological factors and factors in driving performance. The second step of the analysis was step-wise regression using features selected previously as independent variables and change in general driving performance as the dependent variable. As a third step, leave-one-out cross-validation was conducted on the model created with step-wise regression to establish the most accurate coefficients and parameters of the model.

To analyse monitoring of the driver attention during the automation, three types of models were tested. Linear model with the physiological recordings from Measurement 2 as independent variables and number of red cars detected as the dependent variable. Linear model with physiological recordings from the automation period as independent variables and the number of red cars detected as the dependent variable. And the binomial model with the physiological recordings from 30 seconds period before the red car as independent variables and a binary variable indicating if the red car was detected (0-not detected, 1-detected) as a dependent variable. The same approach was used for take-over monitoring, as the analysis consisted of the same three steps. Features selection using uniphysiological linear and binomial regression models, step-wise regression using selected features, and leave-one-out cross-validation were used to establish the most accurate coefficients and parameters of the models. The code for linear and binary leave-one-out cross-validation can be found on the author's GitHub account (https://github.com/SylwiaKaduk). The statistical analysis and data processing were conducted in Matlab R2020a and Excel.

10.2 TAKE-OVER MONITORING

This chapter described an investigation of the possibility to predict driving performance after take-over based on the psychophysiological measurement conducted just before take-over. The measurement was either 2 minutes long recording of the resting state physiology, questionnaires collected before take-over, saliva samples collected before take-over or a voice recording collected before take-over.

Every person has a different physiological baseline as well as different driving abilities. To reduce the confounding effect of the individual differences, regression models used the change in psychophysiology as independent variables and change in general driving performance as the dependent variable. The change in physiology was calculated with a subtraction between the physiological variable from M1 and M3. The change in driving performance was calculated with a subtraction between driving performance during T1 and T2.

A high number of variables created a risk of overfitting and hence there was a requirement for feature selection and reduction. The strategy of features selection and data modelling was graphically represented in figure 10.2. Firstly, each physiological function was used separately to fit a linear regression model using the change in the factors derived from this physiological function recording as independent variables and change in the general driving performance as a dependent variable. Among the statistically significant models, the one with the highest adjusted r-squared was chosen, and significant predictors from this model were selected for further analysis. Secondly, the Bonferroni corrected rho-Spearman correlation was calculated between the factors derived from the physiological recording and all the factors in driving performance. Physiological variables that absolute value of significant rho correlation was equal to or higher than 0.35 were selected for further analysis. As a third step, the rho-Spearman correlation was calculated between the selected physiological variables to avoid collinearity in the model. If the absolute value of rho was equal to or higher than 0.65, the variables were reduced. After such a process of features selection, a multiphysiological model was created using stepwise regression with adjusted r-squared and p values as criteria of the model selection. The model used the change in selected physiological variables as independent variables and change in the general driving performance as a dependent variable. The best selected model was then tested with leave-one-out cross-validation to most accurately estimate the parameters of the model. The following subsections described the process of features selection and the final model creation and testing.

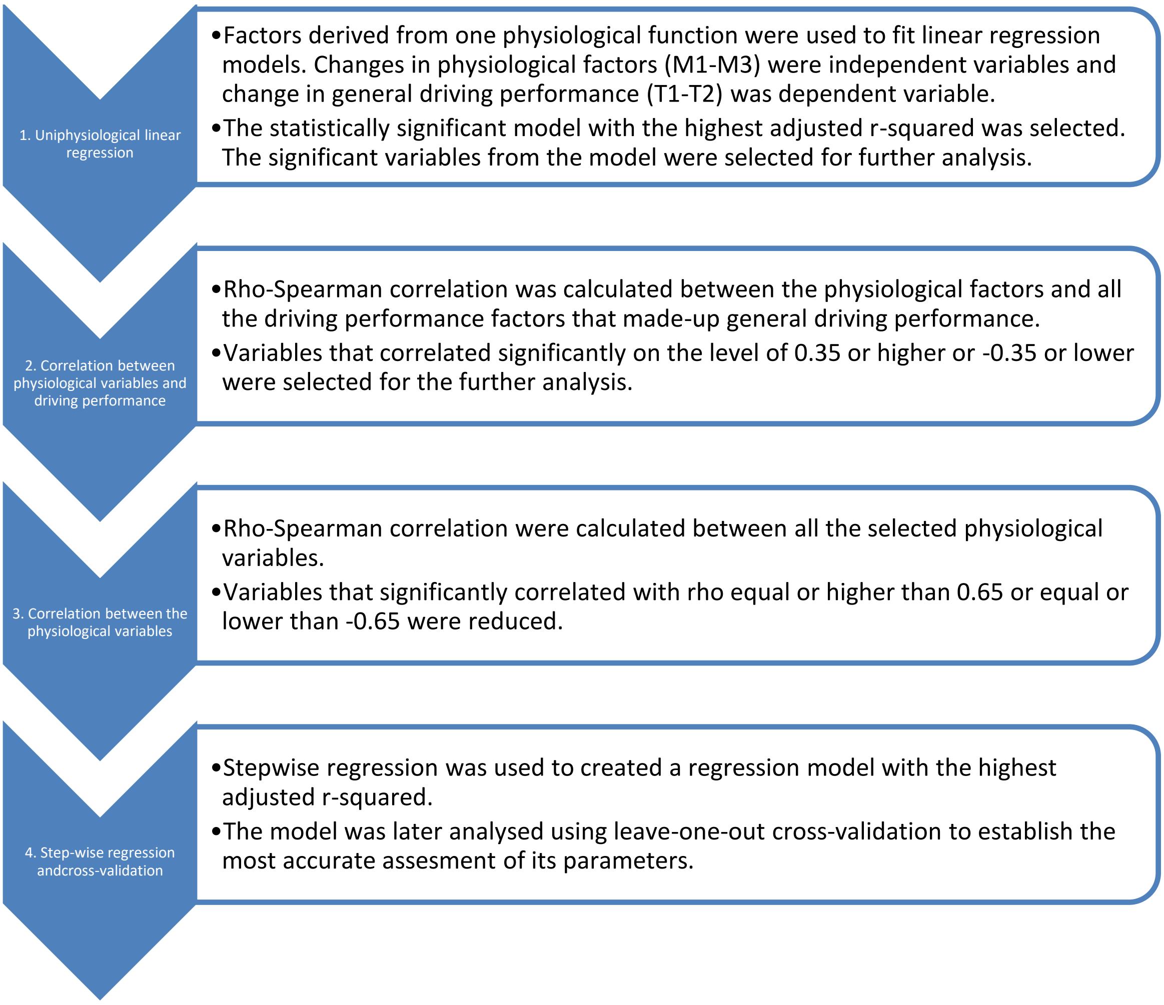

Figure 10.2: Steps of the features reduction and data analysis to create a model predicting driving performance after the take-over based on the physiology of the driver.

10.2.1 QUESTIONNAIRES

To see if the change in results of questionnaires can predict the change in the driving performance, three linear models were tested. The depended variable was the difference between the general driving performance during T1 and T2 (positive result suggested an increase in driving performance and negative a decrease). Initially, the independent variables were a circadian phase, chronotype, and differences between all the questionnaires results measured at M1 (before T1) and M3 (before T2). Later different linear models were tested using the most significant variables from the previous models. All the models can be seen in Appendix 5. The strongest significant model (model 2) used the change in KSS, Fatigue, NASA-TLX Mental Demand, and NASA-TLX Effort scales as independent variables ($F(5,76) = 2.52$, *adjusted r-squared* $=.07$, $p<.05$). The predictive power of the model was very low.

Another analysis was Spearman correlation between the results of the questionnaire before the driving task (M1 before T1 and M3 before T2) and the factors of driving performance following the questionnaires' completion. The table with all the significant correlations can be seen in Appendix 6. The majority of the correlations were low; however, some reached medium values, namely correlation between NASA-TLX frustration scale and standard deviation of the vehicle heading angle (*rs(265) = .34, P<.05*), the correlation between overall NASA-TLX and standard deviation of the steering wheel angle (*rs(265) = .34, P<.05*), and the correlation between overall NASA-TLX and standard deviation of the vehicle heading angle (*rs(265) = .35, P<.05*).

10.2.2 ELECTROCARDIOGRAPHY

Three models were tested to find the best predictive variables in ECG. All the tested models can be seen in Appendix 5. The strongest significant model (model 2) used the circadian phase, change in very high frequencies in heart rate variability, change in sympathetic to vagal tonus ration, and change in RSA as independent variables. Its predictive power was low (*F(5,91) = 3.67 , adjusted r-squared =.1, p<.05*).

As the following step rho-Spearman correlation was calculated between each of the ECG variable and different factors of driving performance during the driving test following the ECG measurement. The values of the significant correlations are presented in Appendix 6. All the significant correlations with ECG variables were small.

10.2.3 VOICE

Three linear models were tested to find the best predictive variables in acoustic voice properties. The details of the models can be seen in the Tables in Appendix 5. None of the models reached statistical significance.

As the following step rho-Spearman correlation was calculated between each of the Voice variables and different factors of driving performance during the driving test following the Voice measurement. The values of the significant correlations were presented in Appendix 6. Some of the correlations' values reached the level of medium association, namely correlation between mean pitch and standard deviation of steering wheel angle (*rs(299) = .37, P<.05*), mean pitch and standard deviation o longitudinal speed (*rs(299) = .30, P<.05*), mean pitch and general driving performance (*rs(299) = -.31, P<.05*), number of pulses and speed exceedances (*rs(299) = -.30, P<.05*), number of pulses and speeding tickets (*rs(299) = -.31, P<.05*), number of pulses and standard deviation of steering wheel angle (*rs(299) = -.39, P<.05*), number of pulses and standard deviation of longitudinal speed (*rs(299) = -.31, P<.05*), number of pulses and general driving performance (*rs(299) = -.31, P<.05*), number of periods and speeding tickets (*rs(299) = -.30, P<.05*), number of periods and standard deviation of longitudinal speed (*rs(299) = -.30, P<.05*), number of periods and general driving performance (*rs(299) = -.33, P<.05*), and number of breaks and general driving performance (*rs(299) = -.30, P<.05*).

10.2.4 ELECTROOCULOGRAPHY

Two linear models were tested to find the best predictive variables in EOG. None of the models reached statistical significance. The details of the models can be seen in the Tables in Appendix 5.

As the following step rho-Spearman correlation was calculated between each of the EOG variables and different factors of driving performance during the driving test following the EOG measurement. The values of the significant correlations were presented in Appendix 6. Almost all of the significant correlations had values on the medium level of association, namely correlation between blinking rate and the number of collisions (*rs(145) = .37, P<.05*), blinking rate and a number of centreline crossings (*rs(145) = .35, P<.05*), the correlation between blinking rate and standard deviation of longitudinal acceleration (*rs(145) = .31, P<.05*), and PERCLOS and number of centreline crossings (*rs(145) = .34, P<.05*).

10.2.5 ELECTRODERMAL ACTIVITY

Two linear models were tested to find the best predictive variables in EDA. The details of the models can be seen in Appendix 5. The strongest significant model (model 2) used the change in SCL Frequency and change in SCL Mean as independent variables. The predictive power of the model was low (*F(5,91) = 3.67, adjusted r-squared =.1, p<.05*).

As the following step rho-Spearman correlation was calculated between each of the EDA variables and different factors of driving performance during the driving test following the EDA measurement. None of the correlations was statistically significant.

10.2.6 ELECTROMYOGRAPHY

Two linear models were tested to find the best predictive variables in EMG. The details of the models can be seen in Appendix 5. None of the models reached statistical significance.

As the following step rho-Spearman correlation was calculated between each of the EMG variables and different factors of driving performance during the driving test following the EMG measurement. The only significant correlation was between the number of collisions and median frequency in corrugator supercilii (*rs(147) = -.21, P<.05*) but its value was low.

10.2.7 RESPIRATION

One linear model was tested to find the best predictive variables in Respiration. It used a change in general driving performance as the dependent variable and change in respiration rate, change in the standard deviation of the respiration, session, and circadian phase as independent variables. It came out to be insignificant and with a very low predictive power. Also, none of the predictors reached statistical significance. As so, no further models were tested. The details of the model can be seen in Appendix 5.

As the following step rho-Spearman correlation was calculated between each of the Respiration variables and different factors of driving performance during the driving test following the Respiration measurement. The significant correlations can be seen in the table in Appendix 6. All of the correlations' values were low.

10.2.8 SALIVA

One linear model was tested to find the best predictive variables in saliva. It used the change in general driving performance as the dependent variable and change in salivary hormonal content, session, and circadian phase as independent variables. It came out to be insignificant and with a very low predictive power. Also, none of the predictors reached statistical significance. As so, no further models were tested. The details of the model can be seen in Appendix 5.

As the following step, the rho-Spearman correlation was calculated between salivary hormonal content variables and different factors of driving performance during the driving test following the saliva collection. The only significant correlation was between cortisol content and stop signs violations; however, it was low. The details of the correlation can be seen in Appendix 6.

10.2.9 OXIMETRY

One linear model was tested to find the best predictive variables in Oximetry. It used the change in general driving performance as the dependent variable and change in Oximetry, session and circadian phase as independent variables. It came out to be insignificant and with a very low predictive power. Also, none of the predictors reached statistical significance. As so, no further models were tested. The details of the model can be seen in Appendix 5.

As the following step, the rho-Spearman correlation was calculated between each of the Oximetry variables and different factors of driving performance during the driving test following the Oximetry measurement. Two correlations were statistically significant; however, their values were low. The details of the correlations can be seen in Appendix 6.

10.2.10 MULTIPHYSIOLOGICAL MODEL

FEATURES SELECTION

Based on the uniphysiological regression models and correlation tables following features were selected for further analysis: change in NASA-TLX scores Mental Demand Scale, change in NASA-TLX overall scores, change in power in very high frequencies in heart rate variability, change in heart rate variability RSA, change in frequency of skin conductance level, change in mean skin conductance level, change in mean pitch of the voice, change in the number of pulses of the voice, change in the number of periods of the voice, change in blinking rate. The analysis of the rho-Spearman correlation between these features demonstrated a significant correlation between change in Change in NASA-TLX Mental Demand scale and change in general scores of NASA-TLX ($rs(265) = .99, P<.05$), change in power in very high frequencies in heart rate variability and change RSA in heart rate variability ($rs(200) = .76, P<.05$), change in very high frequencies in heart rate variability and change in blinking rate ($rs(145) = .45, P<.05$), change in mean pitch of the voice and change in the number of pulses in the voice ($rs(299) = .67, P<.05$), change in mean pitch of the voice and change in the number of periods in the voice ($rs(299) = .59, P<.05$), and change in the number of pulses in the voice and change in the number of periods in the voice ($rs(299) = .93, P<.05$). It allowed to further reduce the number of variables. Final variables selected for the stepwise regression modelling were change in NASA-TLX overall scores, change in power in very high frequencies in heart rate variability, change in mean pitch of the voice, change in the number of periods of the voice, change in blinking rate, change in frequency of the skin conductance level, and change in mean skin conductance level.

STEPWISE REGRESSION

The criteria for stepwise regression features addition or removal were adjusted r-squared and p value. The initial model contained intercept and linear terms for all the predictors. The final model used the change in very high frequency in heart rate variability, change in mean skin conductance level, and the product of the change in overall NASA-TLX and change in frequency of skin conductance level. It had a predictive power of 0.21. The details of the model can be seen in table 10.1.

Table 10.1: Multiphysiolocial model obtained with a step-wise regression modelling method. The independent variables were the most powerful predictors from the selected features and the independent variable was change on the driving performance: Number of observations: 81, Error degrees of freedom: 75, Root Mean Squared Error: 150, R-squared: .26, Adjusted R-Squared: .21, F-statistic vs. constant model: 5.2, p value<0.05

	Estimate	**SE**	**t-Stat**	**P Value**
Intercept	-28.94	27.97	-1.04	NS
Change in power in very high frequencies in HRV	0.01	0.01	2.68	<.05
Change in mean SCL	10.58	5.53	1.91	NS
Change in NASA-TLX*Change in frequency of SCL	-24314	7742.6	-3.14	<.05

CROSS-VALIDATION

Cross-validation was conducted in a leave-one-out way. A regression model with independent variables established with stepwise regression was fit to each training dataset and then tested on values from one experimental session left-out. Statistics for coefficients were calculated as the mean of all the estimates established with leave-one-out cross-validation methods; the same root mean squared error, adjusted r-squared and p value. The mean error of estimation constituted 12.91% of mean driving performance and on average, 184.93% of the observed variable. The details of the final model can be seen in the Table. The comparison of observed values of change in driving performance versus values predicted by the model can be seen in the figure. The association between predictive values and decrease or increase of driving performance was presented in the figure, while the figure showed an association between predictive values and a level of decrease or an increase in driving performance.

Table 10.2: Regression model obtained with a leave-one-out cross-validation using features selected during step-wise regression: Number of observations: 81, Error degrees of freedom: 75, Root Mean Squared Error: 152.25, Adjusted R-Squared: .22, F-statistic vs. constant model: 5.2, p value<0.05, p value<.05

	Estimate	**SE**	**t-Stat**	**P Value**
Intercept	-46.70	19.13	-2.44	<.05
Change in power in very high frequencies in HRV	0.01	0.01	2.62	<.05
Change in mean SCL	11.70	5.30	2.21	<.05

	Estimate	SE	t-Stat	P Value
Change in NASA-TLX*Change in frequency of SCL	-24107.57	5885.41	-4.10	<.05

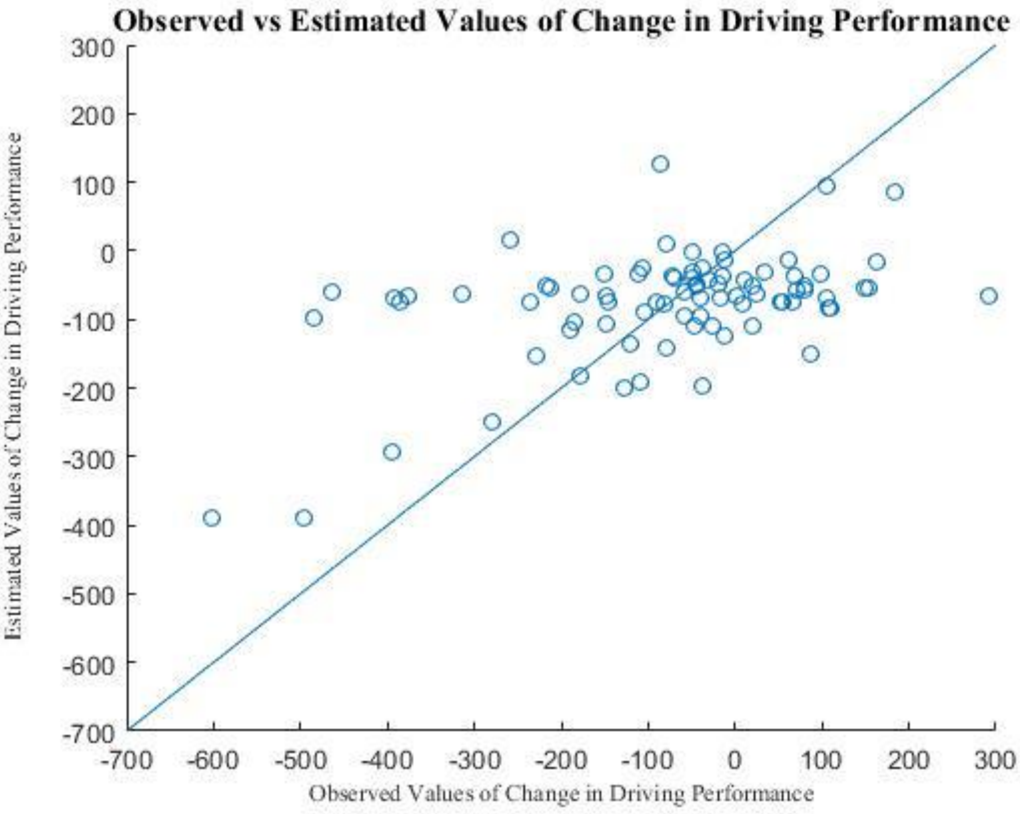

Figure 10.3: Comparison of estimated (y axis) vs observed (x axis) values of change in driving performance. The estimated values were obtained with the final model after cross-validation. The blue line represents a perfect fit of values. The positive values indicate the increased quality of driving performance and negative values a decrease.

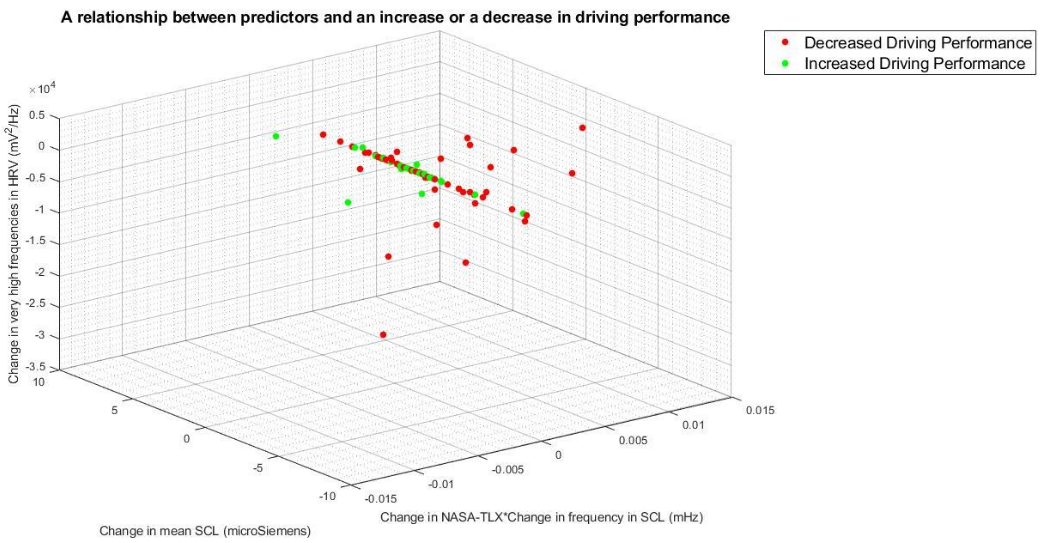

Figure 10.4: A graphical representation of the relationship between the predictors from the model and an increase or a decrease of the driving performance. The red dots represent the decrease and green dots the increase. The negative values of the predictors indicate an increase of the physiological variable and the

positive values a decrease.

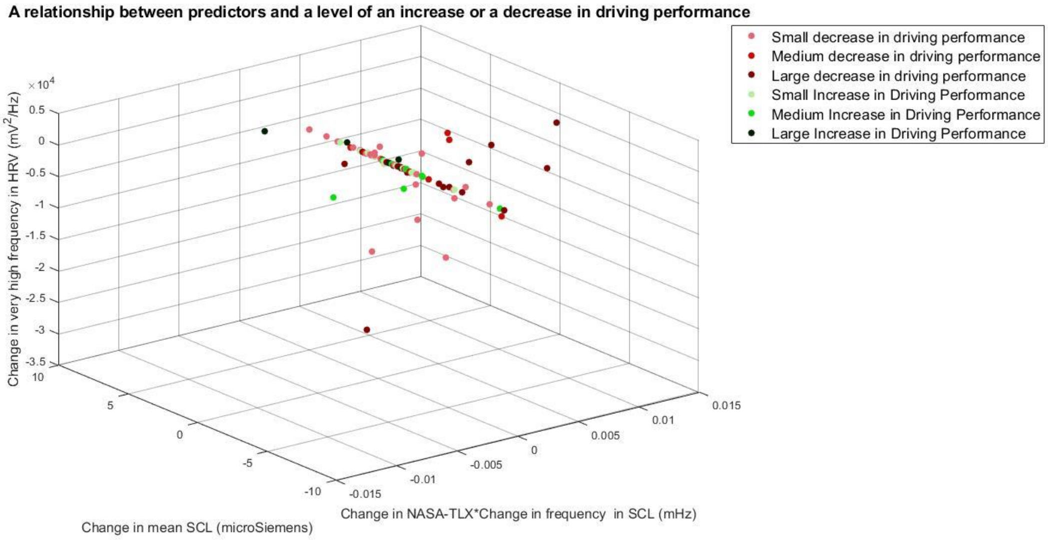

Figure 10.5: A graphical representation of the relationship between the predictors from the model and a level of an increase or a decrease of the driving performance. The different shades of the red dots represent the decrease and various shades of green dots the increase. The negative values of the predictors indicate an increase of the physiological variable and the positive values a decrease.

10.2.11 DISCUSSION

This subchapter presented results of the predictive modelling on driving performance after take-over. The type of take-over analysed was planned; one that could happen when leaving a designated automation driving road, rather than the emergency take-over that could be requested at any time. Some of the results could be extended to the problem of emergency take-over; however, this is beyond the scope of this thesis.

The predictive models presented in this chapter predicted an actual decrease in driving performance rather than the risky driver states that could lead to it. It gave the model a high ecological validity as such a prediction would have an immediate on-road result. Also, the dependent variable in the model was the change in driving performance rather than an absolute value of driving performance. It enabled the prediction of a reduction in driving performance without the confounding effect of the individual driving style or capabilities. The predictive variables included change in physiology, which was a way to avoid the confounding effect of the interindividual physiological differences. That way, the model was more robust regarding interindividual differences as well as presenting ecological validity.

The experimental data consisted of a high number of physiological measurements. To avoid overfitting of the model, there has been a feature selection process applied based on the and linear models. The final model used the following independent variables: change in power in very high frequencies in heart rate variability, change in mean skin conductance level, and a product of the change in NASA-TLX score and change in frequency of skin conductance level.

The mean error of the estimation of the model constituted 12.91% of the mean value of change in driving performance. The model explained 22% of the variance of change in driving performance. Such a value of the coefficient of determination would be treated as low. However, it needs to be taken into account that the model predicted an actual driver performance that is subject to many influences. Also, the physiological functions used to predict the change in the driving performance were not uniquely related to one mental state but to several states that could be both psychological and physiology (such as increased movement). The physiology was recorded during two minutes long resting state to decrease strict bodily influence; however, it still cannot be excluded entirely. As such, this model could not be implemented in road vehicles yet. However, the presented association between psychophysiological functions and driving performance is a good basis for future research. It is also important to note that the future inclusion of EEG signals in the model might significantly increase its predictive power. The use of the different indicators of driving performance could also change the predictive power of the model.

The final predictive model requires some commentary to explain the meaning of the physiological functions included.

The dependent variable was change in general driving performance. The general driving performance was a variable positively correlated with a number of driving mistakes. The higher values of the variable indicated the worse driving performance. The change in driving performance was calculated as a difference between the general driving performance during Driving Test 1 (T1) and Driving Test 2 (T2). The positive values of the change indicated better driving performance during T2 and the negative values indicated worse driving performance during T2.

The first predictive variable was the change in power in very high frequencies in heart rate variability. The decrease in power between Measurement 1 (M1) and Measurement 3 (M3) was associated with a better driving performance during T2. The average value of change in this variable was -411.31 s^2, which indicated an increase in power in very high frequencies in heart rate variability of 411.31 s^2. According to a model, such change would result in a decrease in driving performance of 41.13, which constituted 3.4% of a mean change in driving performance. Previous literature reported that a decrease in power in very high frequencies in heat rate variability was associated 1to the increased mental workload (Brookhuis & De Waard, 2010; Roscoe, 1992; Veltman & Gaillard, 1996; Wilson, 2002)

The second predictive variable was the mean skin conductance level. The decrease in mean SCL between Measurement 1 (M1) and Measurement 3 (M3) was associated with a better driving performance during T2. The average value of change in this variable was -1.26 microSiemens, which indicated an increase in mean skin conductance level of 1.26 microSiemens. According to a model, such change would result in a decrease in driving performance of 14.74, which constituted 1.25% of a mean change in driving performance. Previous literature reported that a reduction in skin conductance level was associated with a decreased mental workload or a decreased arousal (Averty et al., 2002).

The third predictive variable was a product of general NASA-TLX scores and frequency in skin conductance level. The simultaneous increase or decrease in both of these variables was associated with the decreased driving performance during T2, while the reduction of one of them simultaneously with the rise in another was associated with an increase of driving performance during T2. The average value of this product variable was 0.0002. According to a model, such change would result in a decrease in driving performance of 4.82, which constituted 0.41% of a mean change in driving performance. Increased scores in NASA-TLX indicated increased mental workload. Increased frequency in SLC was reported to indicate increased vigilance and arousal (Posada-Quintero et al., 2017). The model would suggest that increased mental workload simultaneous with increased vigilance was associated with a worse driving performance the same as decreased mental workload concurrent with reduced vigilance. At the same time, high mental workload with low vigilance, and low mental workload with high vigilance would result in better driving performance.

All the variables in the final model were previously reported to be associated with mental workload or an arousal level. However, the first variable (power in very high frequencies in HRV) indicated that increased mental workload was associated with improved driving performance. At the same time, the second variable (mean SCL) indicated that a decrease in arousal or mental workload were also associated with the increased driving performance. The third variable suggested that increased mental workload together with increased vigilance was associated with a decreased driving performance, the same as reduced mental workload together with reduced vigilance. In contrast, when one of the variables decreased while the other one increased, then it was associated with improved driving performance. The possible explanation would be related to an ‘inverted U’ shaped association between mental workload and performance and balance between underload and overload (Young & Stanton, 2005). The strongest of the predictors indicates an adverse effect of the underload on the driving performance. The second predictors indicated a negative impact of overload on driving performance. While, the third and the weakest of the predictors, might be related to the balance between arousal and mental workload. The high mental workload with high arousal could lead to more driving mistakes due to hyperactivity and reduced concentration. The low mental workload and low arousal could be related to drowsiness. The high mental workload with low arousal could be associated with the calm state of relaxed concentration, and low mental workload with high arousal to the high cognitive resources. Supporting such an explanation would require additional experimental work and modelling of the association between mental workload, arousal and performance.

10.3 MONITORING OF THE DRIVER STATE DURING AUTOMATION

This chapter described an investigation of the possibility to predict or monitor driver attention during automated mode based on the psychophysiological measurements conducted just before automated mode, during the whole time-course of automated mode or during 30 seconds before attention task (red car detection).

The same as in the case of take-over monitoring the first step of the analysis was features selection for the multiphysiological model. As the first step, three kinds of predictive models were tested for each psychophysiological function separately, linear models to predict attention during automation based on the psychophysiological data gathered before the automated mode, linear models to predict attention during automation based on the physiological data during the whole period of the automation and the binomial models predicting detection of the red car (0- not detected, 1- detected) based on the physiology recorded during 30 seconds before the red car appeared in the simulation.

The linear models investigating the prediction of attention based on the measurement just before the automated mode used two minutes recording of the resting state of the physiology just before the automated mode, questionnaires collected just before automated mode, or voice recorded just before the automated mode. The level of attention was established based on the number of red cars detected during the automation. Such models would allow establishing if the driver can maintain a prolonged, attentive state during automation based on the short measurement before the automated mode.

The linear models investigating the establishment of the level of attention state based on the overall physiological recordings during the automated mode used continuous physiological recordings from the period of thirty-four minutes of the automated mode.

Such a model could bring more knowledge about the general association between sustained attention and physiology over more extended periods.

The binomial models investigating the prediction of the attention test (red car detection) based on the physiological recordings from the 30-seconds period before the red car appeared, used continuous recordings and oximetry recordings from the period of 30-seconds before the red car appeared in the simulation. Such models could display the highest practical utility as they could allow to state if the driver is paying attention at the very moment of the physiological measurement.

A high number of variables created a risk of overfitting and hence there was a requirement for feature selection and reduction. The strategy of features selection and data modelling was graphically represented in figure 10.6. Firstly, each physiological function was used separately to fit a linear or binomial regression model using factors derived from this physiological function recording as independent variables and the number of red cars detected as the dependent variable. In the case of binomial models, the dependent variable was either 0 for the unsuccessful attention test or 1 for the successful attention test. Among the statistically significant models, the one with the highest adjusted r-squared was chosen, and significant predictors from this model were selected for further analysis. As a second step, the rho-Spearman correlation was calculated between the selected physiological variables to avoid collinearity in the model. If the absolute value of rho was equal to or higher than 0.65 variables were reduced. After such a process of features selection, a multiphysiological model was created using stepwise regression with adjusted r-squared and p values as criteria of the model selection. The model used selected physiological variables as independent variables and attention tests as the dependent variable. The best selected model was then tested with leave-one-out cross-validation to most accurately estimate the parameters of the model. The following subchapters described the process of features selection and the final model creation and testing.

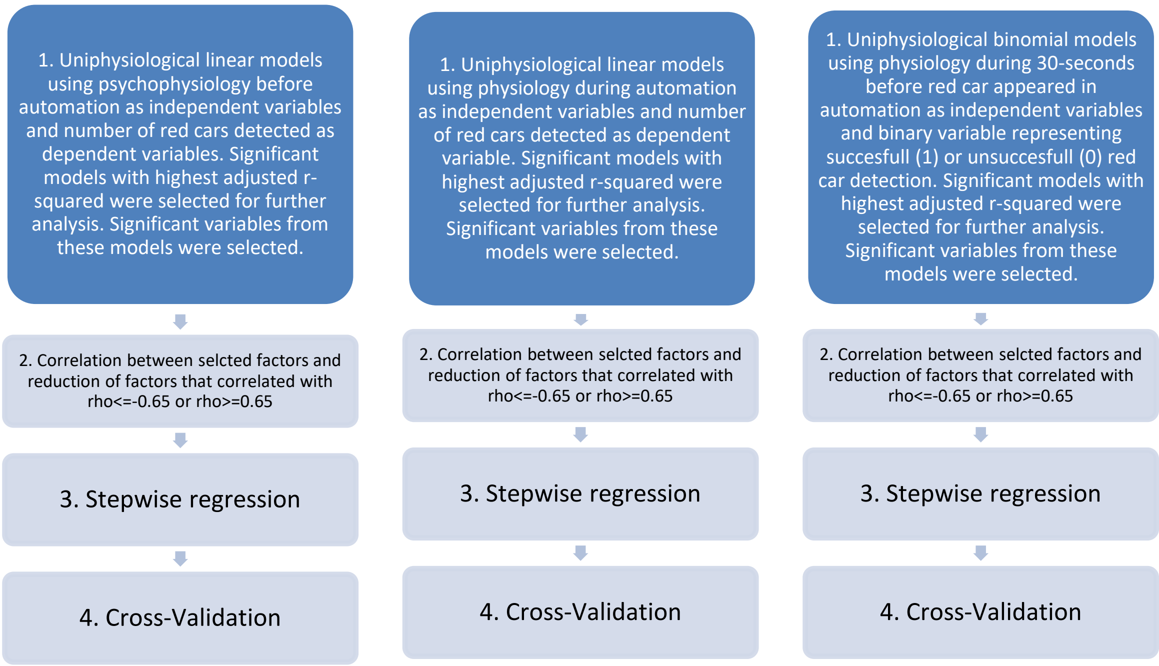

Figure 10.6: Graphical representation of the process of features selection and modelling of psychophysiological monitoring of the driver attention during the automated mode.

10.3.1 QUESTIONNAIRES

To investigate if the number of red cars detected can be predicted with questionnaires filled-out before the automated phase linear regression models were tested. The details of the tested models can be seen in the Tables in Appendix 7. The strongest significant model (model 3) used the circadian phase and fatigue as predictors. The predictive value of the model was very low (*F(3,88) = 3.46, adjusted r-squared =.05, p<.05*).

10.3.2 ELECTROCARDIOGRAPHY

Two linear models were tested to investigate the possibility to predict attention during automation based on the resting state ECG measurement just before the automated mode. They used the number of red cars detected during automated mode as a dependent variable. The details of the models can be seen in the Tables in Appendix 7. Neither of the models was statistically significant.

Two linear models were tested to investigate the association between ECG recording from the whole period of automated mode and the number of red cars detected during the automated mode. The details of the models can be seen in the Tables in Appendix 8. Neither of the models was statistically significant.

Three binomial models were tested to investigate the possibility to predict if participant detected red car (0-did not detect, 1-detected) based on the ECG recording from 30 seconds before the red car appeared in the simulation of automation. The details of the models can be seen in the Tables in Appendix 9. The strongest model (model 2) used heart rate (*Beta = 0.03, t(505) = 1.77, p: NS*), power in very high frequencies in heart rate variability (*Beta = 0.00, t(505) = 1.91, p: NS*), and RSA (*Beta = -0.42, t(505) = -2.23, p<.05*) as independent variables. The predictive power of the model was very low with (*adjusted r-squared = .05*).

10.3.3 VOICE

Two linear models were tested to investigate the possibility to predict attention during automation with acoustic properties of voice collected just before the automated mode. The dependent variable was the number of red cars detected during the automated mode. The details of the models can be seen in the Tables in Appendix 7. Neither of the models was statistically significant.

10.3.4 ELECTROOCULOGRAPHY

Four linear models were tested to investigate the possibility to predict attention during automation with ocular behaviours measured with EOG during a two-minutes resting state just before the take-over. The dependent variable was the number of red cars detected during the following automated mode. The details of the models can be seen in the Tables in Appendix 7. The strongest significant model (model 4) used only PERCLOS as an independent variable. The predictive power of the model was low (*F(2,43) = 14.1, adjusted r-squared =.25, p<.05*).

As a second step, three linear models were tested to investigate the association between ocular behaviours recorded with EOG during the whole automation period and attention during automation measured with the number of red cars detected during the automated mode. The details of the models can be seen in the Tables in Appendix 8. The strongest significant model (model 1) used the circadian phase, blink rate, mean blink duration, PERCLOS, rate of horizontal eye-movements and mean duration of horizontal eye-movements as independent variables. The predictive power of the model was low (*F(7,38) = 3.7, adjusted r-squared =.27, p<.05*).

As a third step, two binomial models were tested to investigate the possibility to predict detection of the red car during automated mode (0-not detected, 1-detected) based on the ocular behaviour measured with EOG during 30 seconds period before the red car appeared in the simulation. The details of the models can be seen in the Tables in Appendix 9. The strongest of the models (model 2) used PERCLOS (*Beta = -0.07, t(225) = -1.15, p: NS*) and rate of horizontal eye-movements (*Beta = 0.05, t(225) = 1.18, p: NS*) as independent variables and a predictive value of the model was middle (*adjusted r-squared=0.35*); however, all of the predictors were insignificant.

10.3.5 ELECTRODERMAL ACTIVITY

One linear model was tested to investigate the possibility to predict attention during automation with EDA variables recorded during the two-minutes long resting-state period just before automation. The dependent variable was the number of red cars detected during automation. The details of the model can be seen in Table in Appendix 7. The model was insignificant with a very low predictive power. Also, all the predictors were insignificant.

As the second step, two linear models were tested to investigate the association between the attention during automation and EDA variables recorded during the whole period of automation. The dependent variable was the number of red cars detected during the automated mode. The details of the model can be seen in the Tables in Appendix 8. Neither of the models was statistically significant.

As a third step, two binomial models were tested to investigate the possibility to predict if the red car was detected (1) or not detected (0) based on the 30 seconds EDA recording just before the red car appeared in the simulation. The details of the models can be seen in the Tables in Appendix 9. The strongest model (model 2) used the mean skin conductance level (*Beta = 0.07, t(460) = 2.36, p<.05*) as a predictor. The predictive power of the model was very low (*adjusted r-squared = 0.01*).

10.3.6 ELECTROMYOGRAPHY

Three linear models were tested to investigate the possibility to predict attention during automation with EMG variables recorded during a two-minutes resting state just before the automated mode. The details of the models can be seen in the Tables in Appendix 7. Neither of the models was statistically significant.

As a second step, two linear models were tested to investigate the association between attention during automation and EMG recorded during the whole automated mode. The dependent variable was the number of red cars detected during automation. The details of the models can be seen in the Tables in Appendix 8. The strongest significant model

(model 2) used the mean frequency of frontalis muscles as an independent variable. The predictive value of the model was (*$F(2,40) = 6.38$, adjusted r-squared =.12, $p<.05$*).

As a third step, two binomial models were tested to investigate the possibility to predict detection of the red car (0-not detected, 1-detected) during the automated mode based on the 30-seconds periods of EMG recording before the red car appeared in the simulation. The details of the models can be seen in the Tables in Appendix 9. Model 1 used the following predictive factors, median frequency of frontalis muscle (*Beta = -0.00, $t(250) = -0.28$, p:NS*), mean frequency of the frontalis muscle (*Beta = 0.01, $t(250) = 1.93$, $p<.05$*), peak frequency of the frontalis muscle (*Beta = 0.01, $t(250) = 0.46$, p:NS*), mean power of the frontalis muscle (*Beta = 0, t(250) = NA, p:NA*), total power of the frontalis muscle (*Beta = 671363.12, $t(250) = 0.65$, p:NS*), median frequency of the corrugator supercilii (*Beta = 0.01, $t(250) = 0.59$, p:NS*), mean frequency of the corrugator supercilii (*Beta = -0.01, $t(250) = 1.08$, p:NS*), peak frequency of the corrugator supercilii (*Beta = 0.00, $t(250) = 0.13$, p:NS*), mean power of the corrugator supercilii (*Beta = 0, t(250) = NA, p:NA*), total power of the corrugator supercilii (*Beta = 6971138.35, $t(250) = 1.76$, p:NS*). It had high predictive power (*adjusted r-squared = 0.73*); however, it was there was the probability that it was over-parametrized and the majority of the predictors were insignificant. Model 2 used Mean Frequency of Frontalis Muscle (*Beta = 0.01, $t(250) = 3.12$, $p<.05$*), and Total Power of Corrugator Supercilii Muscle (*Beta = 3393865.24, $t(250) = 1.97$, $p<.05$*), as independent variables. Both predictors were significant. The predictive power of the model was very low (*adjusted r-squared=0.06*). As so, the high predictive power of Model 1 should be treated as a result of overfitting, and Model 2 should be taken as a realistic description of the predictive power of EMG over red car detection.

10.3.7 RESPIRATION

One linear model was tested to investigate the possibility to predict attention during automation with respiration variables recorded during resting state just before the automated mode. The dependent variable was the number of red cars detected during the automated mode. The details of the model can be seen in Table in Appendix 7. The model was insignificant, with a very low predictive power and all the predictors were insignificant.

Two linear models were tested to investigate the association between respiration variables recorded during the whole automation period and attention during automation. The dependent variable was the number of red cars detected during automation. The details of the models can be seen in the Tables in Appendix 8. The strongest significant model

(model 2) used the respiration rate as an independent variable. The predictive power of the model was very low (*F(2,91) = 6.12, adjusted r-squared =.05, p<.05*).

As a third step, two binomial models were tested to investigate the possibility to predict detection of the red car (0-not detected, 1-detected) during the automated mode based on the 30-seconds periods of EMG recording before the red car appeared in the simulation. The details of the models can be seen in the Tables in Appendix 9. The strongest model (model 2) used the mean breathing rate as a predictor (*Beta = 0.12, t(500) = 2.80, p<.05*). The predictive power of the model was low (*adjusted r-squared = .03*).

10.3.8 OXIMETRY

To investigate the possibility to predict attention during automation with Oximetry based variables linear model predicting the number of red cars detected was tested.

Linear model with Oximetry variables measured during the resting state just before automated mode, circadian phase, session, and chronotype as independent variables and number of red cars detected as the dependent variable was tested. The model was insignificant and displayed a very low predictive power; also, all the predictors were insignificant. The details of the model can be seen in the Table in Appendix 7.

As a further investigation, two binomial models were tested to investigate the possibility to predict if the red car was detected (1) or not detected (0) based on the 30 seconds Oximetry recording just before this red car. The strongest model used pulse as an independent variable (*Beta = 0.03, t(515) = 2.59, p<.05*). The adjusted r-squared was low (*adjusted r-squared = .04*).

10.3.9 MULTIPHYSIOLOGICAL MODELS

FEATURES SELECTION

Based on the uniphysiological models, several variables were selected for further analysis. For the linear model predicting attention during automation based on the physiological recordings conducted just before automation selected variables were fatigue questionnaire scores and PERCLOS.

For the linear model predicting attention during automation based on the whole physiological recording during automation following variables were selected blinking rate, mean blink duration, PERCLOS, rate of horizontal eye-movements, mean duration of horizontal eye-movements, mean frequency of frontalis muscle and respiration rate.

Based on the binomial model predicting the outcome of the attention test based on the 30-seconds physiological recording conducted before the red car appeared in the simulation, the following variables were selected RSA in heart rate variability, mean skin conductance level, mean frequency in frontalis muscle, the total power of corrugator supercilii muscle, mean breathing rate and pulse.

The rho-Spearman correlation was conducted between all these factors to reduce collinearity in the models. Predictors from the model using measurements conducted just before the automation did not significantly correlate with each other.

Among the predictors from the model using measurements conducted during the whole automated mode following variables significantly correlated with each other: respiration rate and rate of horizontal eye-movements ($rs(145) = .31, P<.05$), blink rate and mean blinking duration ($rs(145) = -.55, P<.05$), mean blink duration and mean horizontal eye-movement duration ($rs(145) = .37, P<.05$), PERCLOS and rate of horizontal eye-movements ($rs(145) = .43, P<.05$), blinking rate and PERCLOS ($rs(145) = .61, P<.05$).

Among the predictors from the binomial model following predictors significantly correlated with each other: pulse and RSA ($rs(200) = -.61, P<.05$), RSA and total power in corrugator supercilii muscle ($rs(147) = .25, P<.05$), RSA and mean frequency on frontalis muscle ($rs(147) = -.20, P<.05$), total power in corrugator supercilii muscle and the mean frequency in frontalis muscle ($rs(147) = -.21, P<.05$).

None of the correlations reached an arbitrary cut-off point for reduction. As so none of the variables was reduced before the step-wise regression analysis.

STEP-WISE REGRESSION

Stepwise regression models were conducted in MatlabR2020a using adjusted r-squared and p values as criteria for the addition or removal of the variables. Variables used for the models were previously selected through the process of uniphysiological regression models and reduction based on the correlation that was described in the previous chapters.

The first multiphysiological model used psychophysiological measurements collected just before the automated mode as independent variables and a number of red cars detected as the dependent variable. Initial variables that were used in the model were scores from a fatigue questionnaire and PERCLOS. The final model can be seen in table 10.3. It used only PERCLOS as an independent variable. Its predictive power was low.

Table 10.3: Multiphysiological model fitted with a stepwise regression using adjusted r-squared and p value as criteria for addition or removal of the variables. The dependent variable was the number of red cars detected during the automated mode. Independent variables were measurements collected just before the automation. Number of observations: 43, Error degrees of freedom: 41, Root Mean Square: 0.91, R-squared: .24, Adjusted R-Squared: .22, F-statistic vs. constant model: 13.1, p value < .05

	Estimate	SE	t-Stat	P Value
Intercept	4.88	0.21	23.03	<.05
PERCLOS	-0.14	0.04	-3.62	<.05

The second multiphysiological model used psychophysiological measurements collected during the whole automated phase as independent variables and a number of red cars detected as the dependent variable. Initial variables that were used in the model were blinking rate, mean blink duration, PERCLOS, rate of horizontal eye-movements, mean duration of horizontal eye-movements, mean frequency of frontalis muscle and respiration rate. The final model can be seen in table 10.4. It used blinking rate, mean blink duration, rate of horizontal eye-movements, and a product of PERCLOS and the mean duration of horizontal eye-movements as independent variables. Its predictive power was on the medium level.

Table 10.4: Multiphysiological model fitted with a stepwise regression using adjusted r-squared and p value as criteria. The dependent variable was the number of red cars detected during the automated mode. Independent variables were measurements collected during the whole period of automation. Number of observations: 44, Error degrees of freedom: 37 Root mean square: 0.85, R-squared: 0.41, Adjusted R-Squared: .32, F-statistic vs. constant model: 4.3, p value<.05

	Estimate	SE	t-Stat	P Value
Intercept	8.80	1.15	7.64	<.05
Blinking Rate	-0.10	0.05	-2.19	<.05
Mean blink duration	-7.15	3.19	-2.24	<.05
Rate of horizontal eye-movements	0.05	0.03	1.96	<.05
PERCLOS*mean duration of horizontal eye-movements	1.16	0.53	2.19	<.05

The third multiphysiological model used psychophysiological measurements collected during the period of 30-seconds before the red car appeared in the simulation of automation as independent variables and a binary outcome of the attention test as a dependent variable (0-red car not detected, 1-red car detected). Initial variables that were used in the model were RSA in heart rate variability, mean skin conductance level, mean frequency of frontalis muscle, total power in corrugator supercilii muscle, mean breathing rate and pulse. The final model can be seen in table 10.5. It used RSA in heart rate variability, mean frequency in frontalis muscle, total power in corrugator supercilii muscle, and respiration rate as independent variables. Its predictive power was low.

Table 10.5: Multiphysiological model fitted with a stepwise binomial regression using adjusted r-squared and p value as criteria. The dependent variable was a binary variable representing successful (1) or unsuccessful (0) attention test. Independent variables were measurements collected during the 30 seconds before the red car appeared in the automation. 186 observations, 181 error degrees of freedom, Dispersion: 1, Chi^2-statistic vs. constant model: 24.3, p value<.05, Adjusted r-squared: .15.

	Estimate	**SE**	**t-Stat**	**P Value**
Intercept	0.00	1.76	0.00	<.05
RSA HRV	-0.40	0.18	-2.20	<.05
Mean Frequency in Frontalis	0.01	0.00	2.25	<.05
Total Power in corrugator supercilii	3354800	1424500	2.36	<.05
Breathing Rate	0.15	0.07	2.31	<.05

CROSS-VALIDATION

Cross-validation was conducted in a leave-one-out manner. Regression models with independent variables established with stepwise regression were fit to each training dataset and then tested on values from one experimental session left-out. Statistics for coefficients were calculated as the mean of all the estimates established with leave-one-out cross-validation methods; the same for root mean squared error, adjusted r-squared and p value. The details of coefficients in models after cross-validation can be seen in tables 10.6, 10.7, and 10.8 below.

For the model using resting-state recording just before automation, the mean error of estimation constituted 15.55% of the mean number of red cars detected during the automated phase and on average 13.69% of the observed variable.

The mean error of estimation in the model using a recording from the whole automated mode constituted 8.11% of the mean number of red cars detected during the automated phase and on average 7.09% of the observed variable.

In the case of the binomial model, the mean error constituted 162.30% of the mean result of the attention test and on average 143.69% of the observed variable.

The second and the third predictive models had negative adjusted r-squared values after the cross-validation, suggesting insignificance of the predictors and no predictive power of the models. The first model had low predictive power.

Figure 10.7 presented an association between PERCLOS measured during the resting state before automation and the number of red cars detected during the automation. Figure 10.8 presented a relationship between the observed number of red cars detected and the number of red cars detected estimated by the model. The plots showed that the model systematically underestimated the highest values of the red cars detected and overestimated the lower.

Table 10.6: Multiphysiological model validated with a leave-one-out cross-validation. The dependent variable was the number of red cars detected during the automated mode. Independent variables were measurements collected just before the automation selected with a stepwise regression described in the previous chapter. Number of observations: 43, Error degrees of freedom: 41, Root Mean Square: 0.69, Adjusted R-Squared: .23, F-statistic vs. constant model: 13.1, p value < .05, P value<.05.

	Estimate	**SE**	**t-Stat**	**P Value**
Intercept	4.91	0.20	24.05	<.05
PERCLOS	-0.14	0.04	-3.74	<.05

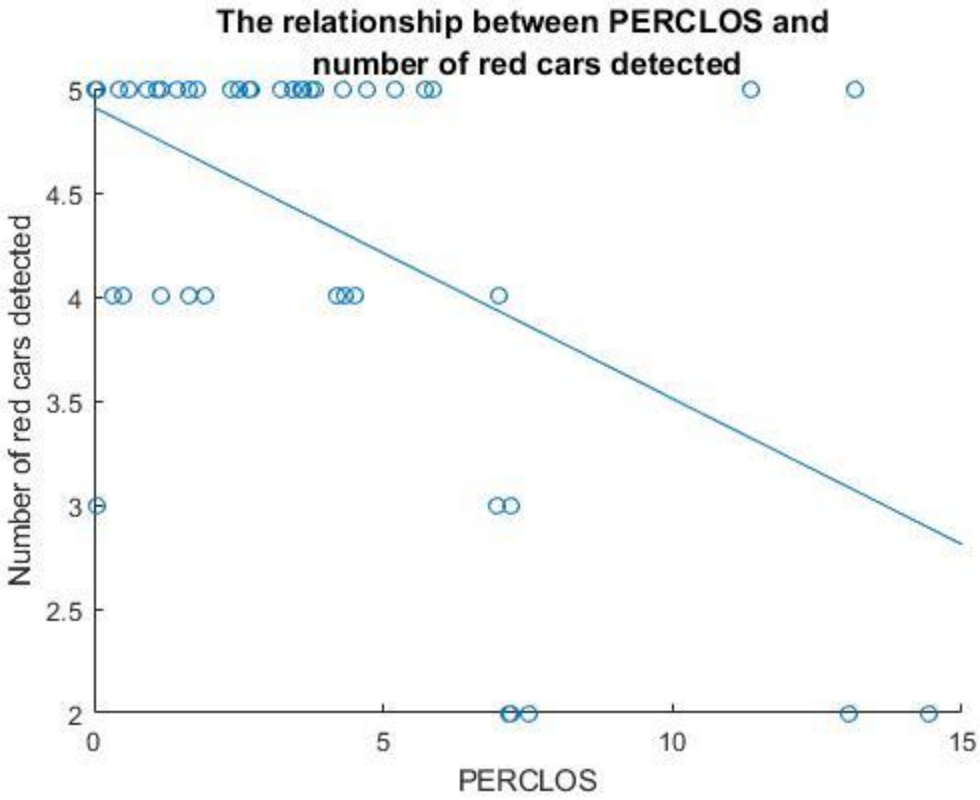

Figure 10.7: Graphical representation of the association between PERCLOS and number of red cars detected. The blue line represents the regression line.

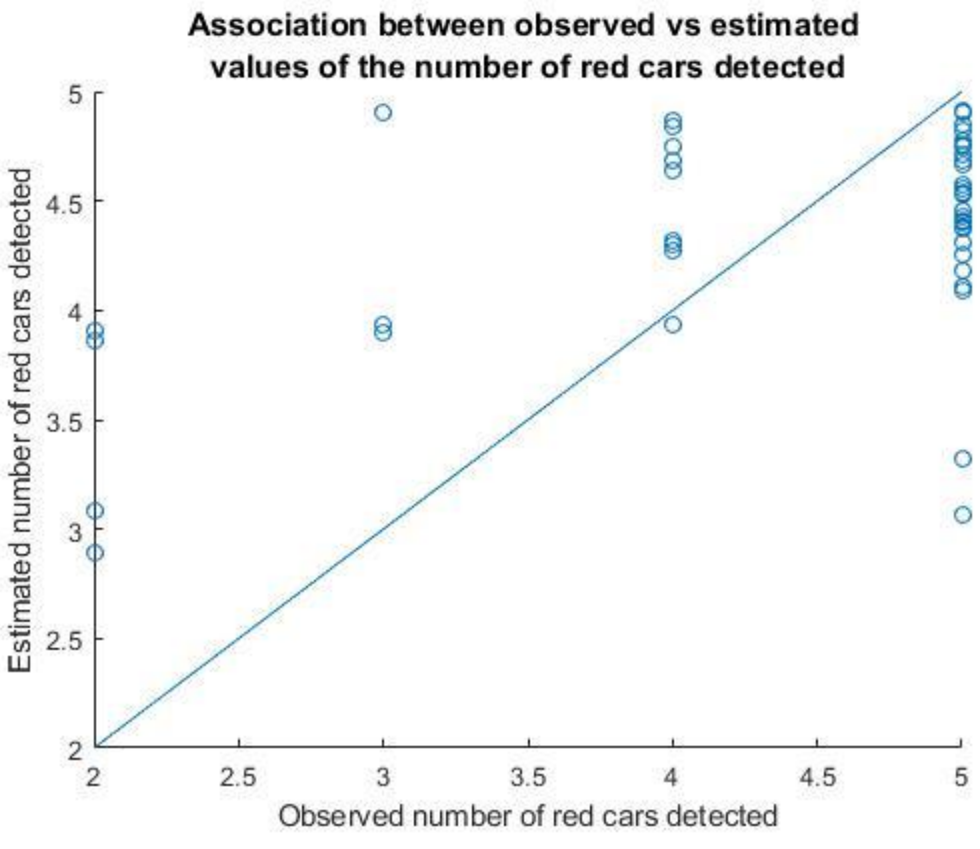

Figure 10.8: The observed values of the number of red cars detected vs. estimated by the model. The blue line represents the 100% fit.

Table 10.7: Multiphysiological model validated with a leave-one-out cross-validation. The dependent variable was the number of red cars detected during the automated mode. Independent variables were measurements collected during the whole automation period selected with a stepwise regression described in the previous chapter. Number of observations: 44, Error degrees of freedom: 37 Root mean square: 0.36, Adjusted R-Squared: -.05, F-statistic vs. constant model: 4.3, p value: NS, adjusted r-squared: -.05

	Estimate	**SE**	**t-Stat**	**P Value**
Intercept	4.08	0.54	7.61	<.05
Blinking Rate	0.01	0.02	0.39	NS
Mean blink duration	1.97	1.81	1.09	NS
Rate of horizontal eye-movements	-0.00	0.02	-0.02	NS
PERCLOS*mean duration of horizontal eye-movements	0.01	0.21	0.07	NS

Table 10.8: Multiphysiological binomial model validated with leave-one-out cross-validation. The dependent variables were selected based on the stepwise regression described in the previous chapter. The dependent variable was a binary variable representing successful (1) or unsuccessful (0) attention test. Independent variables were measurements collected during the 30 seconds before the red car appeared in the automation. 186 observations, 181 error degrees of freedom, Dispersion: 1, Chi^2-statistic vs. constant model: 24.3, p value<.05, Adjusted r-squared: -.01, mean error of estimation: 1.44.

	Estimate	**SE**	**t-Stat**	**P Value**
Intercept	0.34	1.59	0.21	NS
RSA HRV	-0.33	0.16	-2.04	<.05
Mean Frequency in Frontalis	0.01	0.00	2.11	<.05
Total Power in corrugator supercilii	2933397.07	1330408.58	2.21	<.05
Breathing Rate	0.13	0.06	2.10	<.05

10.3.10 DISCUSSION

This subsection investigated a possibility to predict or monitor attention during the automated mode with a variety of psychophysiological measurements. It compared three possible strategies for monitoring. The first one was based on the short psychophysiological measurement conducted before the automated mode. The idea of such a measurement would be to predict if the driver’s state makes them able to keep sustained attention during the following period.

The second approach investigated the association of the physiology recorded during the automation with the measurements of sustained attention.

The third approach aimed to predict attention in the period just before the attentive task.

The automated scenario used in the experiment did not require the driver to intervene at any moment. Participants were asked to stay as attentive as possible and monitor the automated driving process. However, in the absence of required interventions, the task required mainly sustained attention. It was measured with an additional task. Participants were asked to press a button every time they have seen a red car. There were five red cars during the thirty-four minutes of simulation. Such a low number of stimuli created an environment that could easily cause boredom, sleepiness, or distraction. These are the states that might occur in actual automation if it did not require frequent interventions.

The first step of the analysis was features selection based on uniphysiological linear and binomial regression models. Selected features were then reduced based on the collinearity. The resulting variables were used in the process of stepwise regression to create linear and binomial predictive models. As the last step, leave-one-out cross-validation was used to validate the parameters and coefficients of the models. Only one of the models displayed a non-negative adjusted r-squared after the cross-validation.

The final model used PERCLOS measured during the resting state before automated mode as a predictor of the number of red cars detected. The model explained 23% of the variance in the number of red cars detected. The mean error of estimation constituted 15.55% of the mean value of the number of red cars detected. Adjusted r-squared of 0.23 is considered as low; however, it presents a predictive value of just one variable over the sustained attention during the following thirty-four minutes.

The model showed that the increased PERCLOS was associated with further worse sustained attention. PERCLOS was defined as a proportion of time when eyes were closed over a certain period (Abe et al., 2014). In the previous research, increased PERCLOS was often associated with drowsiness (Papadelis et al., 2007; Rodríguez-Ibáñez et al., 2011) and fatigue (Rodríguez-Ibáñez et al., 2011).

10.4 DISCUSSION OF THE DRIVER STATE MONITORING BEFORE TAKE-OVER AND DURING AUTOMATION

This chapter investigated the possibility to monitor driver state in semi-automated driving with different psychophysiological measures. Semi-automated driving is a technological development that could increase driving safety and improve traffic congestion (Kyriakidis et al., 2019); however, it also introduced particular challenges related to human factors. People repetitively fail to maintain sustained attention and often make mistakes when performing monitoring role without many tasks or stimulation (Warm et al., 2008). The analysis presented in chapter 8 suggested that manual driving performance after the take-over might be worse than before the automation. Also, the driver might experience increased sleepiness, fatigue and lower readiness to take-over after automation, as well as decreased mental workload.

Semi-automated driving requires the driver to shift into more monitoring than an active role. Also, the transition from automation to manual driving might create certain risks as the driver might be in a state compromising driving safety. Automation might induce sleepiness, fatigue or distraction (Kyriakidis et al., 2019). A potential solution to these problems might be a system of driver state monitoring.

This chapter described the process, results and analysis of the experimental work investigating a wide variety of psychophysiological measures as methods of driver state monitoring in semi-automated vehicles. The measures used during the experiment were EEG, EOG, EMG, ECG, respiration measurement, EDA, acoustic voice analysis, salivary cortisol and alpha-amylase analysis, pulse, blood oxygenation, Karolinska Sleepiness Scale, Samn-Perelli Fatigue Scale, NASA-TLX, and TORS (take-over readiness scale created for this study). However, EEG analysis was not described in this work due to the time restraints of the doctoral project and the technical problems.

The psychophysiological recordings were analysed concerning their accuracy in the prediction of the driving performance after take-over based on the short measurements before take-over. While their accuracy in prediction of attention during automated mode was based on the short recording just before the automated mode, accuracy in prediction of attention during automated mode based on the continuous recording during the whole automation, and accuracy in prediction of the single result of the attention test during automation based on the 30-seconds recording collected just before the test.

Two models displayed statistical significance and sufficient predictive power. A model predicting driving performance after take-over explained 22% of the variance in the driving performance. It used power in very high frequencies in heart rate variability, mean skin conductance level, and a product of the mean frequency of skin conductance and general scores of NASA-TLX questionnaires. The physiological variables were recorded during the two-minutes resting state just before take-over followed by the collection of the NASA-TLX questionnaire. All of the physiological variables from the model were previously reported to be associated with the mental workload and arousal (Averty et al., 2002; Brookhuis & De Waard, 2010; Posada-Quintero et al., 2017). They suggested a need for an optimal mental workload and optimal arousal level. At the same time, the strongest of the predictors suggests an adverse effect of the underload on the driving performance after take-over. It is possible that the previous automated mode might increase the risk of too low a mental workload.

The second model predicted sustained attention during automated mode with PERCLOS calculated based on the two minutes EOG recording of the resting state just before the automation. It explained 23% of the variance in the attention task during automation. Previous research associated PERCLOS with fatigue and drowsiness (Papadelis et al., 2007; Rodríguez-Ibáñez et al., 2011). The model suggested that increased fatigue and drowsiness before the automation might have the most detrimental effect on sustained attention during the automation and that the ocular behaviours are the most accurate measures that can be used. Both the KSS and fatigues scale proved to be less effective.

The critical point is that subjective readiness to take-over manual driving did not predict well the quality of the manual driving after take-over. Similarly, ocular measurement related to sleepiness and fatigue came out to be more effective in the prediction of attention than the questionnaires. It is consistent with many studies suggesting that drivers are not good in the assessment of their own fitness to the driver (Abe et al., 2014; Schleicher et al., 2008). It underlined the importance of the development of the driver state monitoring methods irrespective of the subjective feeling of the driver.

Both models displayed a low predictive power; however, they were statistically significant and predicted real-life performance in the highly ecological experimental environment. They provided valid information about the role of different physiological functions in driving performance and their possible value as parts of the monitoring systems. Nevertheless, they require more work before implementing into the vehicles.

The inclusion of the EEG recorded during the experiment might improve the models in the future and should be analysed. The use of non-linear models, machine learning or deep learning, could improve the predictive power. Also, these models were trained to predict a particular set of factors related to driving performance. It might be that the use of different measures of driving performance or attention could change their predictive power. For example, STISIM 3 driving simulator did not allow detection of ‘near-misses’ that could be a better indicator of driving safety, than the number of collisions. Not all the measures were used continuously over the whole experiment and it might be that some of them would display higher predictive power if used at more time-points. For example, the voice could be recorded during the automated phase. To achieve high-quality data recording, EOG was used in half of the participants. Increasing the number of participants could improve the predictive power of the models.

It needs to be borne in mind, that the analysed models used physiology to predict an actual driving performance rather than secondary measurements of the driver state. It created a more challenging task, as many physiological variables are associated with multiple psychological and physiological states (Ross et al., 2014) as well as a risky driver states do not always cause a decrease in the driving performance (Parasuraman et al., 2008; Young & Stanton, 2002).

The analysis allowed to support the hypotheses and sub-hypotheses to a certain level. Physiology displayed a distinct possibility to predict driving performance and attention during automation. Also, the physiological states used in the strongest predictive models were previously associated with the risky driver states, namely suboptimal mental workload, sleepiness, and fatigue (Brookhuis & De Waard, 2010; Papadelis et al., 2007; Rodríguez-Ibáñez et al., 2011; Roscoe, 1992).

10.5 CONCLUSIONS

Physiological monitoring has proven to have predictive potential over the driving performance after take-over and attention during automation. However, the models created in this work do not have sufficient predictive power to be implemented in the vehicles. The performance of the models could be improved in a number of ways. The inclusion of EEG recordings, even though, cumbersome could highly increase their predictive power. Machine learning algorithms could select stronger predictive factors from the physiological signal. Inclusion of the driving performance measures together with physiology would also be an option. As PERCLOS displayed predictive power over attention during automation, the use of different ocular measures, like eye-tracking could improve its performance due to the higher accuracy of the measurement and classification.

An optimal level of mental workload and arousal was showed to be the most critical state for driving safety after take-over. At the same time, sleepiness and fatigue came out to more detrimental to the ability of the driver to safely monitor the automation. Because of that interfaces and tasks in the semi-automated vehicles should be designed in a way to improve the level of mental workload and counteract drowsiness and fatigue. Also, more research effort should be invested into a precise understanding of the most optimal balance between mental workload and arousal in a semi-automated environment.

Participants were not effective in prediction of their own driving performance, sleepiness, or fatigue. As so, work on the systems of driver state monitoring seems to be an essential part of the improvement of driving safety. Physiological recording of the driver presents a high challenge for engineers, physiologists and analysts. Current models require more work to have a higher predictive power in real-life driving. Also, physiological sensors need to be developed into less invasive but also more noise-resistant devices. However, driving safety is an important aim raised by governments, scientists, and World Health Organisation (*Ministers to Agree New Global Road Safety Agenda to 2030*, n.d.; World Health Organization, 2015) and the opportunity to improve it should not be neglected.

11. CONCLUSION

11.1 INTRODUCTION

This work described doctoral research into driver state monitoring in semi-automated vehicles from the circadian perspective. The main aim of this work was to evaluate and compare different methods of driver state monitoring in semi-automated vehicle. The comparison was aimed to select the psychophysiological measures and develop a model that would be the most effective in the prediction of actual driver performance in the semi-automated vehicle. However, the experiment assumed a perspective of the circadian rhythmicity and also allowed the analysis of the patterns in driver state throughout the time-course of the semi-automated driving.

The literature reviews identified risky driver states that could jeopardize safety in semi-automated driving and the methods of psychophysiological monitoring that have the potential to detect these states. Three following literature reviews identified circadian effect on manual driving and suggested a possibility of the circadian effect on the semi-automated driving and driver state monitoring. It was hypothesised that semi-automated driving might be riskier at night and during the so-called ‘mid-afternoon’ dip and that circadian phase, which should be taken into account when interpreting a signal from the monitoring devices.

The experiment investigated the time-course of the semi-automated driving during the day and at night to predict manual performance after the take-over as well as attention during semi-automated mode using psychophysiological measures. An additional chapter provided recommendations for the laboratory construction and set-up for multiple physiological recordings, giving an example of the faraday cage constructed for this experimental work.

This work presented several notable findings that were listed in Table 11.1. A summary of these findings is presented in the next section.

Table 11.1: Notable findings presented in this thesis with corresponding chapters.

Theme	**Notable finding**	**Corresponding Chapters**
The decrease in Driving Performance and Mental Underload Related to Automation	Driving Performance decrease after automation	8
	Sleepiness and fatigue increase after automation	8
	Decrease of readiness to take-over after automation	8

Theme	Notable finding	Corresponding Chapters
	Decrease of mental workload after automation	8
	Mental Underload after automation	8
	Drivers are not accurate in the assessment of own fitness to drive	8, 10
Circadian Effect	Circadian effect on semi-automated driving	5, 6, 8, 9
Monitoring of the Driver State for Take-Over Safety	**The predictive model of driving performance after take-over**	**10**
	Suboptimal mental workload and arousal are the most detrimental for performance after take-over	**10**
Monitoring of Attention During Automation	**The predictive model of attention during automation**	**10**
	Sleepiness and fatigue are most detrimental for attention after automation	**10**
Technical Recommendations for the Laboratories Recording Multiple Physiological Functions Simultaneously	Recommendations for high-quality laboratory construction and set-up for multiple physiological recordings	7

11.2 SUMMARY OF FINDINGS

11.2.1 DECREASE IN DRIVING PERFORMANCE AND MENTAL UNDERLOAD RELATED TO AUTOMATION

The main focus of the analysis on the time-course of semi-automated driving was the change between the measurement before and after the automated mode. Such a change could be related to the effect of automation on the driver; however, due to the experimental design, it could not be dissociated from the effect of time-on-task. The analysis showed a decrease in manual driving performance after automation, increase of sleepiness, increase of fatigue, decrease of readiness to take-over manual control over the vehicle, decrease of mental workload, increase of power in low frequencies in heart rate variability, a decrease of the number of pulses in voice, decrease of the number of periods in voice, and a decrease in pulse. These results allowed to support hypotheses 3 and 4 that, Driver psychophysiological state and performance differ before and after automated mode, and that driver performance is worse after automated mode, while their psychophysiological state is related to the lower cognitive state.

The decrease in driving performance was more generalized at night. It allowed to partially support hypothesis 5 that driver psychophysiological state and performance in semi-automated vehicles differ between day and night, and sub-hypothesis that, driving performance and attention during automation decrease during the night.

These results could suggest underload during automated driving resulting in poorer subsequent manual performance after the take-over, as well as reduced readiness to drive and feeling more tired. It could imply that automation might negatively influence manual driving safety due to the underload (Young & Stanton, 2002, 2002). It is of particular concern because a period of automated driving is likely to be followed by manual driving if automated driving is reserved for motorway use only. This effect is pronounced for night-time driving. Additionally, as participants were not accurate in the assessment of their own fitness to drive there is a need for a supplementary assessment of the driver state independent from their self-awareness.

11.2.2 CIRCADIAN EFFECT

This work presented the analysis of the association between the circadian phase and psychophysiological variables from the day and night sessions recorded during the resting state and driving performance. It allowed supporting hypothesis 5 that driver psychophysiological state and performance in semi-automated vehicles differ between day and night.

The analysis showed a variety of differences in the performance and physiology; however, the direction of the effect for the driving performance was not consistent, which did not allow to support the sub-hypothesis that: driving performance and attention during automation decrease during the night. Previous research showed that drivers performed worse at night only when the circadian phase was in the interaction with the sleep deprivation (Matthews, Ferguson, Zhou, Kosmadopoulos, et al., 2012; Matthews, Ferguson, Zhou, Sargent, et al., 2012). Analysis of stepwise regression using a number of hours slept before the experiment, circadian phase, and KSS results did not support it. However, this experiment did not use all the circadian phases and the participants were rarely sleep deprived even during the night experiment.

Also, when the time-course of automation was analysed separately for day and night sessions, the decrease of driving performance after automation was more generalized at night. It was described in chapter 8. It partially supported the sub-hypothesis that: driving performance and attention during automation decrease during the night. Participants tended to feel sleepier, more fatigued, and less ready to take-over at night. EOG and EMG displayed circadian fluctuations and they were previously used in the systems of the driver state monitoring. This supported previous predictions based on the literature review that the circadian phase might affect the interpretation of the measures in the driver state monitoring systems.

11.2.3 MONITORING OF THE DRIVER STATE FOR TAKE-OVER SAFETY

The analysis of driver state monitoring aimed to create predictive models that used psychophysiology to predict manual driving performance after take-over. It consisted of a process of features selection, step-wise regression, and leave-one-out cross-validation. The take-over monitoring model explained 22% of the variance in the change of driving performance before and after automation. Predictive variables were change in resting-state power in very high frequencies in heart rate variability, change in resting-state mean skin conductance level, and a product of the change in NASA-TLX score and change in resting-state frequency of skin conductance level. All of the variables have been associated in previous studies of mental workload and arousal (Averty et al., 2002; Brookhuis & De Waard, 2010; Posada-Quintero et al., 2017; Roscoe, 1992). It allowed to partially support hypothesis 1 that psychophysiological measurements of the driver during semi-automated driving can provide a prediction of the driving performance after take-over, and sub-hypothesis that psychophysiological indicators that can predict driving performance are related to one of the risky states identified in chapter 2, namely sleepiness, fatigue, distraction, mental workload or situation awareness.

11.2.4 MONITORING OF ATTENTION DURING AUTOMATION

The second predictive model aimed to predict attention during the automated mode. It explained 23% of the variance in attention during automation with PERCLOS measured during two minutes resting state just before the automated mode. In previous research, PERCLOS was used as an indicator of sleepiness and fatigue (Papadelis et al., 2007; Rodríguez-Ibáñez et al., 2011). It allowed to partially support the hypothesis 2 psychophysiological measurements of the driver during semi-automated driving can provide a prediction of their attention during the automated mode of semi-automated driving, and sub-hypothesis that psychophysiological indicators that can predict driver attention are related to one of the risky states identified in chapter 2, namely sleepiness, fatigue, distraction, mental workload or situation awareness.

11.2.5 TECHNICAL RECOMMENDATIONS FOR THE LABORATORIES RECORDING MULTIPLE PHYSIOLOGICAL FUNCTIONS SIMULTANEOUSLY

Chapter 7 provided a prototype for the support tool for the researchers. It combined the literature into a decision tree helping researchers to select the most optimal laboratory set-up for the multiple simultaneous physiological measurements. It also described various strategies of noise reduction in the physiological signals. The decision tree was a result of the challenging set-up of this experiment. This experiment used over ten psychophysiological measures simultaneously, which is an exception in comparison to the experiments described in the literature. It required a very conscious use of the recording tools and multiple measures against the data noise. The laboratory was constructed to reduce signal noise. The process of laboratory construction was also described in chapter 7.

11.3 EVALUATION OF THE RESEARCH APPROACH

Multiple parts of this research were novel, unique and highly challenging due to the complexity of the experimental set-up and the number of psychophysiological measures that were taken. Multiple physiological signals were recorded simultaneously in the semi-automated set-up. Such experimental designs are often not undertaken because it is highly challenging to collect multiple physiological recordings without loss in the data quality. Due to these issues, there was a diligent preparation of the laboratory and experimental set-up that was described in Chapter 7 (Sweeney et al., 2012).

The recording of the multiple physiological signals simultaneously allowed a comparison of the measures and evaluation of their joined usability. However, at the same time, it decreased the statistical power of the analysis. Hence, it was impossible to conduct more demanding models in terms of numbers of observations, for example, non-linear regression.

The majority of previous investigations into driver state monitoring have validated predictive models with the questionnaire results or alternative physiological recordings. For example, they created the models of monitoring and checked their accuracy based on the sleepiness declared by the participants (Brookhuis & De Waard, 2010; Murata et al., 2015). This research attempted more ecological validity and directly predicted driving performance. It was a highly challenging approach, because performance is a result of multiple factors, while the majority of the physiological functions is not uniquely related to the one mental state but might change in a variety of circumstances (Brookhuis & De Waard, 2010; Papadelis et al., 2007). This research also provided a novel perspective on the circadian effect on semi-automated driving. It created a theoretical basis as well as a partial validation of the theoretical predictions.

This research was not without some weaknesses and they need to be borne in mind when interpreting the results. Also, some parts of the experimental design were inevitable, but different variations of this study in future could bring more knowledge and understanding of the topic. Even though EEG was measured during the experiment, it was not used for the analysis due to technical reasons and temporal constraints. Adding EEG to the predictive models could have increased their power, as well as add some insight into the physiology and cognition of the driver in semi-automated vehicles. Not all the psychophysiological measures were used at all the time-points and not in all participants. For example, the voice was only recorded during three measurement points, while EOG and EMG were recorded only in half of the participants. Considering that one of the predictive models used PERCLOS as an only independent variable, increasing the number of participants could strengthen the model. Analysing the physiology of the participants with low and high driving capability could also increase the depth of analysis. Also, the analysed predictive models were only linear or binomial. At the same time, quadratic or exponential regression could explain changes in the driving performance better, especially considering the non-linear effect of mental workload on performance. To fully dissociate the impact of automation from the effect of time-on-task an experiment simulating semi-automated driving should be compared to a control condition without the automated phase with the same duration. To better understand the impact of the circadian phase on semi-automated driving more circadian phases should be studied in the forced desynchrony study design using the whole 24-hours cycle for the analysis (Dijk et al., 1992; Matthews, Ferguson, Zhou, Sargent, et al., 2012). The long duration of the experiment and a large number of measurement methods led to discomfort in some of the participants. Due to that, some of the participants had to take breaks during the experiment. To reduce the confounding effect of the unscheduled breaks, the most promising methods should be tested in a shorter experiment. Finally, this study was undertaken in a low-fidelity driving simulator. Therefore the predictive models should also be tested in an on-road vehicle. The measures of driving performance and attention used by the simulator could also impact the predictive power of the models. In future studies, other methods could be tested, for example, ‘near-misses’ that were not detected by STISIM 3.

11.4 IMPLICATIONS OF THE RESEARCH

This research showed that semi-automated driving might be risky due to the influence of automated mode on driver state. Participants’ driving was worse, and they felt less alert after automation. Some parts of their physiology suggested mental underload. This effect was more prominent at night; but, when different factors were compared between the circadian phases, the results were not consistent. Also, participants were not able to accurately assess their own fitness to drive other ways of ensuring a safe driver state should be investigated.

A wide variety of psychophysiological functions was investigated as methods of driver state monitoring, which was the main purpose of this research work. Two viable monitoring models were proposed to ensure attention during the automated mode and fitness to drive before the take-over. The models showed a potential of heart rate variability, skin conductance, NASA-TLX questionnaire, and some electrooculography-based factors to provide a prediction of driver performance. However, the predictive power of the models was too low to be implemented in vehicles at this stage. They should be treated as a basis for further research. In a conclusion, physiological measures that are recommended for driver state monitoring are electrocardiography, electrodermal activity, and electrooculography. However, the predictive power of the models proposed in this thesis is not sufficient to immediately apply them in the vehicle.

The interpretation of the psychophysiological factors in the models as well as the psychophysiological changes through the time-course of the semi-automated driving, could suggest a particular association between driver state and driver performance in semi-automated driving. Drowsiness and fatigue seemed to be the most disruptive states for the successful monitoring of the automation while insufficient mental workload and arousal for safe manual driving performance after take-over from vehicle automation. It is also essential to bear in mind that the self-awareness of the driver and their subjective assessment of their own mental workload, sleepiness and fatigue were not sufficient to effectively predict their performance. Therefore there is a requirement for additional monitoring systems.

These results were partially consistent with previous studies and predictions. They confirmed concerns related to the underload during automation and its negative effect on driver performance (Kyriakidis et al., 2019; Young & Stanton, 2002, 2002, 2005). They also confirmed that physiology could be used as a predictor of driver performance (Brookhuis & De Waard, 2010; Sahayadhas et al., 2015); however, the accuracy of predictions does not meet standards for commercial implementation at the moment. The expectation of the negative effect of the circadian phase on driving (Kaduk et al., 2020) was only partially confirmed.

The research presented in this thesis had some unique and novel characteristics. To the knowledge of the author, it is the only experiment studying such a large number of physiological measures in the same driving experiment. Also, the topic of driver state monitoring and driver physiology is quite widely studied, but not in semi-automated driving. Moreover, the majority of the studies that investigated driver state monitoring used physiological measurements to predict driver state measured by other methods, such as questionnaires, secondary tasks or EEG but did not attempt to predict actual driving performance (Borghini et al., 2014; Di Stasi et al., 2011; Rodríguez-Ibáñez et al., 2011). To the knowledge of the author, it is also a first attempt to assess the circadian effect on semi-automated driving. The decision tree with the recommendations for the multiphysiological measurements in the laboratory was also a unique attempt to create guidance for research using many psychophysiological measures simultaneously.

The shortcomings of this doctoral project provided some recommendations for manufacturers, engineers, and researchers. Semi-automated driving seems to decrease driving safety and the automated phase might negatively affect driver state. Until level five of automation is introduced, the human driver should be the main actor controlling vehicles (Banks & Stanton, 2016). In the case of the introduction of semi-automated systems, special care should be taken to avoid sleepiness and suboptimal mental workload. Systems of driver state monitoring are recommended; however, they require more scientific work before implementation in vehicles. Psychophysiological measures that are recommended for further research are electrocardiography, electrodermal activity, and electrooculography.

11.5 FUTURE WORK

This research showed a possibility to predict driver performance in semi-automated driving with psychophysiological measurements. The most recommended psychophysiological measures were electrocardiography, electrodermal activity, and electrooculography. However, the proposed models were not sufficiently advanced to be implemented in road vehicles. Further work on the modelling of the proposed physiological functions is recommended. The models should be investigated in a more realistic driving environment. Also, alternative modelling strategies could bring more predictive power. Quadratic and exponential predictive models could better depict non-linear associations. Machine learning and deep learning could also prove to be useful to understand exact second-by-second physiological changes associated with a decrease of the performance during automation and to dissociate exact physiological factors that could be further used in more accurate predictive models.

The circadian phase should be included in the research about semi-automated driving as it proved to affect the performance and physiology of the driver; however, understanding this effect requires more scientific work. It should be studied in the forced-desynchrony protocol with recording collected during the whole 24-hours cycle.

Special research effort should be invested in the understanding and mathematical modelling of the interaction between mental workload, arousal and driving performance. Linear modelling seems to be insufficient to understand this association, so quadratic models, exponential models, and machine learning and deep learning methods would be recommended.

As automation seems to put the driver in the role of cognitive underload, the future design of the semi-automated systems should place the driver in more active and involving tasks.

11.6 CLOSING REMARKS

Semi-automated driving shifts the role of the driver to the more monitoring role for the automated periods of the driving (Kyriakidis et al., 2019). This research, in-line with previous concerns, suggested that such a situation might negatively affect driver state and decrease driving safety after the take-over. It is, therefore, recommended to keep the driver as the primary active agent in the driving process until level 5 of automation is introduced to the market. Drivers are not accurate in the assessment of their own fitness to drive, so if semi-automated systems were to be introduced to the vehicle or other disciplines of technology, it is recommended to combine them with devices to monitor the state of the driver or other type of operator.

This work proposed a system to predict performance after take-over with heart rate variability, skin conductance level, and NASA-TLX and a system to predict attention during the automated mode with PERCLOS. However, the predictive power of the models is not sufficient to be implemented to the vehicles and requires more scientific work before. The circadian effect proved to influence both performance in semi-automated driving and physiology of the driver; however, to fully understand the association more work is required with the forced-desynchrony research protocol and measurements during the whole 24-hours cycle. The driver states that proved to be most detrimental for the quality of the take-over were suboptimal mental workload and suboptimal arousal. At the same time, drowsiness and fatigue negatively affected attention during the automation.

For future research, it is recommended to investigate the association between driving performance, mental workload, and arousal. To effectively use them in the driver monitoring system, mathematical modelling of the association is required. Quadratic and exponential regression models could be more useful than linear regression. Also, machine learning and deep learning would be recommended to analyse second-by-second physiological changes in driver state during automation to fully understand a performance decrease and the physiological changes during the automation in a precise temporal resolution.

This research showed that semi-automated vehicles are not yet proven to be completely safe, and the current potential of driver monitoring systems are not yet sufficiently advanced to determine if drivers are ready to assume manual control. Full automation is not yet ready to be introduced and semi-automated systems need more work in the area of driver state monitoring to improve safety. Further research is also needed in understanding the second-by-second psychophysiological changes that occur in the operator during the process of automation monitoring. At this stage of knowledge, it is better to keep people more involved in the active controlling processes rather than to shift them towards the monitoring role.

12. APPENDIX 1- LIST OF RISKY DRIVER STATES WITH DEFINITIONS

Table 12.1: The list of risky driver states and their definitions identified during the literature review described in the chapter 2.

Factors	**Frequency of Factors**	**Definition**	**List of articles**
Drowsiness	58	No definition provided	(Akrout & Mahdi, 2016), (Boverie, Giralt, & Le Quellec, 2008), (Grace et al., 1998), (Higgins et al., 2017), (Kumari & Kumar, 2017), (Leng, Giin, & Chung, 2015), (C. C. Liu, Hosking, & Lenné, 2009), (Mittal, Kumar, Dhamija, & Kaur, 2016), (Petridou & Moustaki, 2000), (Popieul, Simon, & Loslever, 2003), (Smith, Shah, & da Vitoria Lobo, 2003)
		No definition, but authors provide psychological, behavioural or physiological symptoms	(Choi & Kim, 2017), (Chuang, Huang, Ko, & Lin, 2015), (Daza et al., 2011), (Ebrahim, Stolzmann, & Yang, 2013), (Fairclough & Graham, 1999), (Ha & Yoo, 2016), (Jackson, Raj, et al., 2016), (He et al., 2017), (Kartsch, Benatti, Rossi, & Benini, 2017), (B.-G. Lee, Jung, & Chung, 2011), (C.-T. Lin et al., 2006), (C.-T. Lin et al., 2008), (C. T. Lin et al., 2010), (Maglione et al., 2014), (Murata, Fujii, & Naitoh, 2015), (Murata, Naitoh, &

Factors	Frequency of Factors	Definition	List of articles
			Karwowski, 2017), (Park, Xu, Sridhar, Chi, & Cauwenberghs, 2011), (Rodríguez-Ibáñez, García-González, Fernández-Chimeno, & Ramos-Castro, 2011), (Sahayadhas, Sundaraj, Murugappan, & Palaniappan, 2015), (Solaz et al., 2016), (Van Winsum, 2000), (Vitabile, De Paola, & Sorbello, 2010), (Vural et al., 2007), (J. Wang, Sun, Fang, Fu, & Stipancic, 2017), (M. Wang et al., 2016), (Y. Wang, Xin, Bai, & Zhao, 2017), (Wierwille, Wreggit, Kirn, Ellsworth, & Fairbanks, 1994), (V. E. Wilkinson et al., 2013)
		State in between wake and sleep	(Johns, 2000), (Jackson, Kennedy, et al., 2016)
		Increased sleep propensity	(Chipman & Jin, 2009), (G. Li, Lee, & Chung, 2015)
		Last period of wakefulness and the initial period of stage-1 sleep	(Kwai, Imtiaz, Bowyer, & Rodriguez-Villegas, 2016), (Sahayadhas, Sundaraj, & Murugappan, 2013), (Thorslund, 2004)
		The first period of stage-1 sleep	(Yeo, Li, Shen, & Wilder-Smith, 2009)
		Psychological and physiological state related to the lack of sleep	(Ma'touq et al., 2014)
		Feeling tired, weary, lacking energy and motivation	(Borghini, Astolfi, Vecchiato, Mattia, & Babiloni, 2014), (Dhupati, Kar, Rajaguru, & Routray, 2010), (Qian et al., 2016), (da Silveira, Kozakevicius, & Rodrigues, 2016), (Yang & Jeong, 2015)
Fatigue	40	No definition	(Häkkänen & Summala, 2001), (Mittal et al., 2016)
		No definition, but authors provide psychological, behavioural or physiological symptoms	(Ahn, Nguyen, Jang, Kim, & Jun, 2016), (Damousis & Tzovaras, 2008), (Deng, Xu, Yang, Miao, & Northeastern Univ, 2010), (Leandro L Di Stasi et al., 2015), (Fan, Sun, Yin, & Guo, 2010), (Fu, Wang, & Zhao, 2016), (Grace et al., 1998), (Haq & Hasan, 2016), (Jap, Lal, Fischer, & Bekiaris, 2009), (Lal & Craig, 2005), (B.-G. Lee et al., 2011), (T. Li et al., 2015), (J. Liu, Zhang, & Zheng, 2010), (Maglione et al., 2014), (J. F. May & Baldwin, 2009), (Min, Wang, & Hu, 2017), (Park et al., 2011), (Popieul et al., 2003),

Factors	Frequency of Factors	Definition	List of articles
			(Qiong, Jingyu, Mingwu, & Yujie, 2006), (Simon et al., 2011), (Smith et al., 2003), (Tran, Thuraisingham, Wijesuriya, Craig, & Nguyen, 2014)
		Performance decrease and subjective feeling of weariness, lack of motivation and unwillingness to work caused by prolonged work or stress	(Borghini et al., 2014), (I. D. Brown, 1994), (Chipman & Jin, 2009), (Milosevic, 1997), (Johns, 2000), (Puspasari, Iridiastadi, & Sutalaksana, 2015)
		State in between wake and sleep caused by prolonged work	(Lal & Craig, 2001), (Lal & Craig, 2002)
		Mode of cognitive and physical energy saving in the high work environment	(G. Matthews & Desmond, 2002)
		The feeling of being tired	(Chakraborty & Aoyon, 2014), (Dhupati et al., 2010), (Heikoop, de Winter, van Arem, & Stanton, 2016), (Melnicuk et al., 2016), (Wijesuriya, Tran, & Craig, 2007)
		Decreased performance	(Bundele & Banerjee, 2009)
Behavioural Distraction	26	No definition, but examples of distracting behaviours provided	(Bando & Nozawa, 2015), (Brookhuis, de Vries, & De Waard, 1991), (Caird, 2015), (Chan, Nyazika, & Singhal, 2016), (Drews, Yazdani, Godfrey, Cooper, & Strayer, 2009), (Haigney, Taylor, & Westerman, 2000), (Horrey & Wickens, 2006), (Hosking, Young, & Regan, 2009), (Klauer et al., 2014), (Lamble, Kauranen, Laakso, & Summala, 1999), (J. D. Lee, Roberts, Hoffman, & Angell, 2012), (Márquez, Cantillo, & Arellana, 2015), (Petridou & Moustaki, 2000), (Redelmeier & Tibshirani, 1997), (Rumschlag et al., 2015), (Seiler, 2015), (Strayer & Drew, 2004),
		Attention directed towards activity competing with the main task, ‘eyes-off-road’	(Hoel, Jaffard, Boujon, & Van Elslande, 2011), (Hosking et al., 2009), (Klauer et al., 2014), (Liang & Lee, 2010), (Regan, Hallett, & Gordon, 2011), (Ross et al., 2014), (Sahayadhas, Sundaraj, Murugappan, & Palaniappan, 2015), (Taib, Yu, Jung, Hess, & Maier, 2013)
Cognitive distraction	24	No definition	(Casner, Hutchins, & Norman, 2016), (Chakraborty & Aoyon, 2014), (Dogan et al., 2017), (Parker, Reason, Manstead, & Stradling, 1995)

Factors	Frequency of Factors	Definition	List of articles
		No definition, but authors provide psychological, behavioural or physiological symptoms	(Bando & Nozawa, 2015), (Heikoop et al., 2016), (Ji & Yang, 2002), (Melnicuk et al., 2016), (Merat, Jamson, Lai, Daly, & Carsten, 2014), (Miyaji, Kawanaka, & Oguri, 2009), (Sahayadhas et al., 2015), (Smith et al., 2003)
		Attention directed towards thoughts not related to the main task, 'mind-off-road'	(Yanchao Dong, Hu, Uchimura, & Murayama, 2011), (He, Becic, Lee, & McCarley, 2011), (Hoel et al., 2011), (Liang & Lee, 2010), (Parnell et al., 2016), (Regan et al., 2011), (Sahayadhas et al., 2015), (Stanton & Salmon, 2009), (Taib et al., 2013), (Wesley, Shastri, & Pavlidis, 2010), (Yang & Jeong, 2015)
		Decreased vigilance	(Kawanaka, Miyaji, Bhuiyan, & Oguri, 2013)
Sleepiness	23	No definition	
		No definition, but authors provide psychological, behavioural or physiological symptoms	(Åkerstedt et al., 2013), (Åkerstedt, Peters, Anund, & Kecklund, 2005), (Daza et al., 2011), (A. J. Filtness, Armstrong, Watson, & Smith, 2017), (Ftouni et al., 2013), (Horne & Baulk, 2004), (Jackson, Raj, et al., 2016), (Krajewski, Batliner, & Golz, 2009), (Krajewski, Schnieder, Sommer, Batliner, & Schuller, 2012), (Lowden, Anund, Kecklund, Peters, & Åkerstedt, 2009), (Maglione et al., 2014), (Murata et al., 2017), (Papadelis et al., 2007), (Perrier et al., 2016), (Resalat, Saba, & Afdideh, 2012), (Y. Wang et al., 2017), (Watling, Armstrong, & Radun, 2015)
		The propensity to fall asleep	(Johns, 2000)
		Last period of wakefulness and an initial period of stage-1 sleep	(Thorslund, 2004)
Suboptimal Mental workload	12	No definition, but authors provide psychological, behavioural or physiological symptoms	(Gregersen & Bjurulf, 1996), (Recarte & Nunes, 2003), (Stanton, Young, & McCaulder, 1997)
		The disproportion between one's mental resources and task demands	(Borghini et al., 2014), (Brookhuis & de Waard, 2010), (Engström, Markkula, Victor, & Merat, 2017), (De Winter, Happee, Martens, & Stanton, 2014), (Heikoop et al., 2016), (Melnicuk et al., 2016), (Palinko, Kun, Shyrokov, & Heeman, 2010), (M. S. Young & Stanton, 2002),

Factors	Frequency of Factors	Definition	List of articles
		Using too many mental resources for the competing task and not having enough for the main task	(Yang & Jeong, 2015)
Insufficient situation awareness	11	No definition	(Borghini et al., 2014), (Casner et al., 2016), (Dogan et al., 2017), (Hancock, 2015), (Kyriakidis et al., 2017), (Merat et al., 2014), (Stanton & Young, 1998)
		'The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future', knowing what's going on and what to do.	(De Winter et al., 2014), (Endsley, 1996), (Heikoop et al., 2016), (Johannsdottir & Herdman, 2010), (Parasuraman, Sheridan, & Wickens, 2008), (Stanton & Salmon, 2009)
Intoxication	11	A factor not included in further analysis	(Stanton & Salmon, 2009), (World Health Organization, 2015), (Bekiaris, 1999), (Jonah, 1986), (Richer & Bergeron, 2009), (Arnedt, Wilde, Munt, & MacLean, 2001), (Petridou & Moustaki, 2000), (Fairclough & Graham, 1999), (Rogeberg & Elvik, 2016), (Sloan, McCutchan, & Eldred, 2017), (Behnood & Mannering, 2017)
Anger	11	A factor not included in further analysis	(Stanton & Salmon, 2009), (Melnicuk et al., 2016), (Jonah, 1986), (Stanton & Young, 1998), (Roidl, Frehse, & Höger, 2014), (Öz, Özkan, & Lajunen, 2010), (Arnett, Offer, & Fine, 1997), (Garrity & Demick, 2001), (Kawanaka et al., 2013), (G. Matthews, 2002), (Minhad, Ali, & Reaz, 2017)
Sleep	6	No definition provided by the analysed papers	(Horne & Reyner, 1995), (Sagberg, 1999), (J. Horne & Reyner, 1999), (Häkkänen & Summala, 2001), (Sahayadhas et al., 2013), (Higgins et al., 2017)
Stress	6	A factor not included in further analysis	(Heikoop et al., 2016), (Stanton & Salmon, 2009), (Damousis & Tzovaras, 2008; Petridou & Moustaki, 2000), (Garrity & Demick, 2001), (G. Matthews, 2002)
Bad health	3	A factor not included in further analysis	(Stanton & Salmon, 2009), (Häkkänen & Summala, 2001), (Vuurman, Vuurman, Lutgens, & Kremer, 2014)
Loss of vigilance	2	A factor not included in further analysis	(Guo, Pan, Zhao, Cao, & Zhang, 2018), (Qiong et al., 2006)

Factors	Frequency of Factors	Definition	List of articles
Haste	1	A factor not included in further analysis	(Stanton & Salmon, 2009)
Motion Sickness	1	A factor not included in further analysis	(Diels & Bos, 2016)
Arousal	1	A factor not included in further analysis	(Heikoop et al., 2016)
Sleep Inertia	1	'grogginess, disorientation, and sleepiness that can accompany awakening from a nap'	(Wörle et al., 2020)

13. APPENDIX 2- LIST OF METHODS OF DRIVER STATE MONITORING

Table 13.1: The list of the driver state monitoring methods with the exact state indicators identified in the literature review presented in the chapter 3.

Method	The frequency of the method in the literature	Measured State	An indicator of the state	References
Electroencephalography (oscillations)	32	Drowsiness	Increase in alpha power	(Dhupati et al., 2010)
			Increase in delta power	(Rodríguez-Ibáñez et al., 2011)
			Increase in the delta, theta and alpha power over occipital areas, a general decrease of beta power	(Borghini et al., 2014)

Method	The frequency of the method in the literature	Measured State	An indicator of the state	References
			Increase in absolute and relative theta and delta power	(Oken et al., 2006)
			Alpha: theta ratio decrease	(Oken et al., 2006)
			Increase in alpha and theta power, decrease in beta power	(Horne & Baulk, 2004; Lal & Craig, 2001)
			Reappearing alpha trains that gradually disappear	(Lal & Craig, 2001)
			Increase in centro-frontal alpha power and a decrease of occipital alpha amplitude in 1-10 seconds series	(Lal & Craig, 2001)
			Increased blink duration, cumulative alpha and alpha lasting longer than 3 seconds	(Kartsch et al., 2017)
			Simple thresholding classifier using Cz frequencies and relative alpha and delta power	(Patrick et al., 2016)
			Malahanois distance analysis with alpha and theta frequencies	(Lin et al., 2008)
			ICA of EEG power spectra	(Chuang et al., 2015; Lin et al., 2006)
			γ/δ, $(\gamma+\beta)/(\delta+\alpha)$ indices decrease	(T. L. da Silveira et al., 2016)
			Bayesian nonnegative CP decomposition	(Qian et al., 2016)
			ANN classification of the wavelet transform	(Subasi, 2005)
			Slow eye movements	(Jiao & Lu, 2016)
			Absolute and relative powers and signal entropy complexity	(Hwang et al., 2016)
			The alpha increase, synchrony increase	(Papadelis et al., 2007)

Method	**The frequency of the method in the literature**	**Measured State**	**An indicator of the state**	**References**
			SVMPPM analysis	(G. Li et al., 2015)
		Vigilance decrease	Increase in theta power	(Borghini et al., 2014; Lal & Craig, 2001)
			Increase in slow frequencies power	(Oken et al., 2006)
		Mental workload	Increase in theta power over frontal, prefrontal, central and parietal cortex	(Borghini et al., 2014)
			Decreased alpha and increased delta power	(Wilson, 2002)
			SVM and Elastic Net on alpha power, theta power and the spectral characteristic	(Hogervorst, Brouwer, & Van Erp, 2014)
			Linear DFA	(Berka et al., 2007)
			B-alert classification system	(Berka et al., 2004)
			Increased lower alpha power	(Kamzanova et al., 2014)
		Fatigue	Increase in theta power	(Borghini et al., 2014; Lal & Craig, 2001)
			Alpha spindles	(Borghini et al., 2014; Simon et al., 2011)
			Increase in alpha and theta activity	(Boksem & Tops, 2008; Lal & Craig, 2001; Perrier et al., 2016)
			Increase in theta and delta	(Lal & Craig, 2002, 2005)
			Increase in alpha power	(Rodríguez-Ibáñez et al., 2011)
			KPCA-HMM and complexity parameters	(J. Liu et al., 2010)
			Deep belief algorithm analysis of single-channel Fp1	(P. Li et al., 2016)
			The increase of the ratio of slow waves to fast waves	(Jap et al., 2009)

Method	The frequency of the method in the literature	Measured State	An indicator of the state	References
			Alpha decrease, theta increase, θ/α increase, (θ+α)/β increase	(Cheng & Hsu, 2011)
		Behavioural distraction	Increase in theta power	(Borghini et al., 2014)
		Sleep	The decrease of alpha power, an increase of theta and delta power	(Oken et al., 2006)
			Increase in theta waves, an appearance of sleep spindles and k-complex	(Lal & Craig, 2001)
			MLP-FF analysis of spectral entropy	(Sriraam et al., 2016)
Hybrid of methods	20	Drowsiness	SVM- EEG and fNIRS	(Ha & Yoo, 2016)
			Increased occipital alpha with slow eye movements	(Yeo et al., 2009)
			Multilevel ordered logit on average eyelid closure, average pupil diameter, SD of the lateral position and steering wheel reversals	(Wang & Xu, 2016)
			ECH, PPG and EDA	(Hwang et al., 2016)
			ANN and Random Forest algorithm on multiple driving performance features and EEG	(Wang et al., 2016)
			KNN analysis of PCA fusion of EMG and ECG	(Sahayadhas et al., 2015)
			SVM model based on EDA and plethysmography	(Leng et al., 2015)
			Analysis of steering wheel movements and pulse	(Sanpeng et al., 2010)
			Combination of oximetry and PERCLOS	(Sharma & Bundele, 2015)
		Cognitive distraction	AdaBoost analysis of pupils size and an interval between heart rate R-waves	(Miyaji et al., 2009)

Method	The frequency of the method in the literature	Measured State	An indicator of the state	References
			AdaBoost analysis of gaze direction, pupils diameters, head orientation and ECG	(Kawanaka et al., 2013a)
		Stress	An algorithm based on ECG, EDA and EMG	(Healey & Picard, 2005)
			SVM, Decision Tree and Naive Bayes Algorithm using EDA, Blood volume pulse, pupil diameter and skin temperature	(Barreto et al., 2007)
			Hidden Markov Model using EEG, EMG and respiration	(Fu et al., 2016)
			EEG, EOG, ECG and fNIRS	(Ahn et al., 2016)
		Mental workload	ANN analysis of EEG, EOG, heart rate and respiration	(Wilson & Russell, 2003)
		Fatigue	Hidden layers MLP NN analysis of oximetry and skin conductance	(Bundele & Banerjee, 2009; Sharma & Bundele, 2015)
Eye-tracking	20	Mental workload	An increase in the pupil size	(Di Stasi et al., 2011; Marinescu et al., 2018; Palinko et al., 2010)
			SVM and Elastic Net on Pupils size, blink rate and blink duration	(Hogervorst, Brouwer, & Van Erp, 2014)
			A decrease of spontaneous eye movements, saccades and saccades extent	(May et al., 1990)
			Increased pupil size and decreased blink duration	(Ahlstrom & Friedman-Berg, 2006)
		Drowsiness	PERCLOS	(Abe et al., 2014; Brookhuis & De Waard, 2010; Grace et al., 1998; Rodríguez-Ibáñez et al., 2011; Wilkinson & Stretton, 1971; Yang & Jeong, 2015)

Method	The frequency of the method in the literature	Measured State	An indicator of the state	References
			Pupils size decrease	(Oken et al., 2006)
			Increased blink duration, delayed lid opening, decreased lid closure speed	(Schleicher et al., 2008)
			Increased pupil diameter, increased eye closure duration, increased blinks duration, increased number of saccades, increased number of off-road fixation, longer off-road fixations, faster saccades	(J. Wang et al., 2017)
			Fuzzy Expert System	(Damousis & Tzovaras, 2008)
			Increased blinking frequency	(He et al., 2017)
		Cognitive distraction	A decrease in pupils size	(Kristjansson et al., 2009)
			PERCLOS	(Rodríguez-Ibáñez et al., 2011)
			A decrease of microsaccade velocity	(Di Stasi et al., 2015)
Electrocardiography	18	Mental workload	A decrease of the 0.1 Hz power in heart rate variability	(Fairclough & Houston, 2004)
			An increase in heart rate and a decrease in heart rate variability	(Brookhuis & De Waard, 2010; Wilson, 2002)
			Increased heart rate	(Averty et al., 2002; Maglione, Borghini, Aricò, et al., 2014; Wilson, 2002)
			A decrease in heart rate variability	(Roscoe, 1992; Veltman & Gaillard, 1996)
			Decrease in inter-beat-interval	(Veltman & Gaillard, 1996)
		Drowsiness	A decrease in heart rate	(Maglione, Borghini, Aricò, et al., 2014; Oken et al., 2006)

Method	The frequency of the method in the literature	Measured State	An indicator of the state	References
			LDA on heart rate variability	(Vicente et al., 2016)
			Analysis of spectral features with KNN classifier	(Sahayadhas et al., 2015)
			A decrease of mean power frequency	(Murata & Hiramatsu, 2008)
		Behavioural distraction	Analysis of spectral features with KNN classifier	(Sahayadhas et al., 2015)
		Fatigue	A decrease in heart rate	(Lal & Craig, 2001, 2002)
		Stress	An increase in heart rate	(Ogorevc et al., 2011; Schreinicke et al., 1990)
		Anger	SVM classification based on the root mean square successive difference and heart rate variability	(Minhad et al., 2017)
Electrooculography	16	Mental workload	Blink rate and blink duration decrease	(Borghini et al., 2014; Richter et al., 1998; Veltman & Gaillard, 1996)
		Drowsiness	Increased blink rate, a decrease of saccadic eye movements	(Borghini et al., 2014)
			Increased slow eye movements	(Oken et al., 2006; Shin et al., 2011)
			Increased blink duration, delayed lid opening, decreased lid closure speed	(Schleicher et al., 2008)
			PERCLOS	(Papadelis et al., 2007; Rodríguez-Ibáñez et al., 2011)
			LDS analysis	(Zhu et al., 2014)
			Increased blink duration, the increased standard deviation of the lateral eyes position	(Ingre et al., 2006)

Method	**The frequency of the method in the literature**	**Measured State**	**An indicator of the state**	**References**
			Increased blink rate	(Papadelis et al., 2007)
			A decrease in spontaneous blink rate and an increase of slow eye movements	(Minhad et al., 2017)
		Sleep	Slow lateral eye-movements	(Oken et al., 2006)
		Fatigue	An increase of blinks speed and disappearance of saccadic movements	(Lal & Craig, 2001, 2002)
			An increase of blink rate	(Stern et al., 1994)
			PERCLOS	(Rodríguez-Ibáñez et al., 2011)
			A decrease in blink amplitude and blinking rate, increase in eye closure time	(Morris & Miller, 1996)
Functional Near-infrared spectroscopy	9	Anger	An increase of DLPFC oxygenation asymmetry	(Aranyi et al., 2015)
		Drowsiness	Oxygenation increase	(Khan et al., 2016)
			LDA with mean oxyhemoglobin, signal peak and sum of peaks	(Khan et al., 2016; Khan & Hong, 2015)
		Fatigue	Oxyhemoglobin increase in the frontal lobe	(Li et al., 2015)
			Oxyhemoglobin decrease over frontal and superior temporal cortices	(Suda et al., 2009)
		Cognitive distraction	Deoxyhemoglobin decrease in the frontal lobe	(Li et al., 2015)
		Mental workload	Increased activation in the prefrontal lobe	(Mehta & Parasuraman, 2013)
			Increased oxygenation	(Ayaz et al., 2012)
			Increased oxygenation in DLPFC	(Bunce et al., 2011)
Electrodermal Activity	9	Mental workload	Increased EDA	(Aranyi et al., 2015; Miyake et

Method	The frequency of the method in the literature	Measured State	An indicator of the state	References
				al., 2009; Wilson, 2002)
			Increased skin conductance	(Averty et al., 2002)
			Increased number of spontaneous fluctuations	(Richter et al., 1998)
		Stress	Increased EDA	(Ogorevc et al., 2011; Perala & Sterling, 2007)
		Drowsiness	Decreased EDA	(Michael et al., 2012)
		Sleep	The gradual decrease of EDA	(Hwang et al., 2015)
Acoustic speech analysis	7	Drowsiness	A decrease in the voiced/unvoiced consonants duration	(Dhupati et al., 2010)
			SVM analysis of features based on unvoiced and voiced consonants duration, LFCC, MFCC, LPC, HNR, LTAS	(Krajewski, Batliner, et al., 2009)
			AdaBoost and Bagging algorithms classifications of different phonetic features of speech	(Krajewski et al., 2012)
			LDA and ANN analysis of prosodic and spectral speech characteristics	(Krajewski & Kröger, 2007)
		Fatigue	Narrowing of spectral range and reduction of speech intensity	(Milosevic, 1997)
			Multiple classification algorithms with various acoustic features	(Krajewski, Trutschel, et al., 2009)
			A decrease of a fundamental frequency	(Whitmore & Fisher, 1996b)
Electroencephalography (event-related potential)	7	Drowsiness	Analysis of SSVEP with SVM and LDA	(Resalat et al., 2012)
		Mental workload	P300 decrease in latency and decrease in amplitude	(Brookhuis & De Waard, 2010)

Method	The frequency of the method in the literature	Measured State	An indicator of the state	References
			SVM and Elastic Net algorithms on Pz ERP	(Hogervorst, Brouwer, & Van Erp, 2014)
		Decreased vigilance	Decreased amplitude of ERPs	(Oken et al., 2006)
		Sleep inertia	Delayed P300	(Bastuji et al., 2003)
			A smaller peak-to-peak amplitude of N1-P2 components	(Kolff et al., 2003)
		Fatigue	P300 latency at Pz decrease	(Cheng & Hsu, 2011)
		Cognitive distraction	N1 negativity and Nb2 component in reaction to irrelevant stimuli larger or the same than in response to relevant stimuli.	(Boksem & Tops, 2008)
Electromyography	6	Sleep	Pharyngeal dilator decrease	(Oken et al., 2006)
		Behavioural distraction	Analysis of spectral features with KNN classifier	(Sahayadhas et al., 2015)
		Anger	Corrugator supercilii, levator palpebrae superiolis, orbicularis oculi increased activity	(Van Boxtel, 2010)
		Drowsiness	Analysis of spectral features with KNN classifier	(Sahayadhas et al., 2015)
		Mental workload	Corrugator supercilii and frontalis increase, jaw decrease	(Cohen et al., 1992)
		Stress	The tension in the trapezius muscle	(Healey & Picard, 2005)
Subjective report	6	Drowsiness	Result in the sleepiness scale	(Ftouni et al., 2013; Horne & Baulk, 2004; Ingre et al., 2006; Kaida et al., 2006; Oken et al., 2006; Van Dongen & Dinges, 2000)
Blood pressure	6	Mental workload	A decrease in blood pressure variability	(Brookhuis & De Waard, 2010)

Method	**The frequency of the method in the literature**	**Measured State**	**An indicator of the state**	**References**
			Increased blood pressure	(Veltman & Gaillard, 1996)
		Fatigue	Increase in diastolic blood pressure	(Milosevic, 1997)
		Stress	Increased blood pressure	(Larkin et al., 1998; Ogorevc et al., 2011; Schreinicke et al., 1990)
Infrared video camera	5	Drowsiness	PERCLOS	(Vitabile et al., 2010)
			Increased blink duration	(Caffier et al., 2003; Ftouni et al., 2013)
		Cognitive distraction	Changes in supraorbital facial temperature	(Wesley et al., 2010)
		Mental workload	The decreased temperature on the tip and sides of the nose	(Marinescu et al., 2018)
Facial expression	4	Fatigue	Rubbing face, yawning, nodding, slow eye-lid closure, decreased facial tonus	(Lal & Craig, 2001)
		Drowsiness	Eye closure, head rotation, yawning	(Smith et al., 2003; Vural et al., 2007)
			Gabor based analysis of the facial features	(Fan et al., 2010)
Saliva analysis	4	Drowsiness	Decreased alpha-amylase	(Pajcin et al., 2017)
		Mental workload	Increased cortisol	(Zeier et al., 1996)
			Increased immunoglobulin	(Zeier et al., 1996)
		Stress	Increased cortisol	(Schreinicke et al., 1990)
			Increase in alpha-amylase	(Perala & Sterling, 2007)
Body temperature	4	Fatigue	Ear body temperature increased	(Milosevic, 1997)
		Mental workload	A decrease of the temperature on the tip of the nose	(Itoh, 2009)
		Drowsiness	Decreased body temperature	(Michael et al., 2012)

Method	**The frequency of the method in the literature**	**Measured State**	**An indicator of the state**	**References**
		Stress	Heat Flux	(Ogorevc et al., 2011)
Pupillometry	3	Mental workload	PCA and ICA pupils size change	(Jainta & Baccino, 2010)
		Drowsiness	A decrease in the pupil size	(Ranzijn & Lack, 1997)
			Pupillographic sleepiness test	(Wilhelm et al., 2015)
Respiration	3	Sleep	Fall in abdominal respiration compared to thoracic	(Oken et al., 2006)
		Drowsiness	Change in breathing frequency stability	(Rodríguez-Ibáñez et al., 2011)
		Fatigue	Change in breathing frequency stability	(Rodríguez-Ibáñez et al., 2011)
		Stress	Increased breathing rate	(Schreinicke et al., 1990)
Driving performance	2	Cognitive distraction	Steering and throttling acceleration	(Bando & Nozawa, 2015)
			Velocity, steering angle, longitudinal acceleration, gas pedal use, brake pedal use, steering wheel angle, lateral acceleration,	(Yang & Jeong, 2015)
		Drowsiness	Velocity, gas pedal use, steering wheel angle	(Yang & Jeong, 2015)
		Mental workload	Velocity, longitudinal acceleration, gas pedal use, brake pedal use, steering wheel angle, lateral acceleration	(Yang & Jeong, 2015)
Psychomotor performance	2	Drowsiness	Decreased speed and accuracy in psychomotor tasks	(Lal & Craig, 2001; Van Dongen & Dinges, 2000)
Body position	2	Drowsiness	Increase in horizontal and vertical neck bending, back pressure on the seat	(Murata et al., 2015, 2017)
Doppler blood flow meter	1	Mental workload	A decrease of the blood flow on the tip of the nose	(Miyake et al., 2009)
Actigraphy	1	Sleep	Actigraphy	(Mullaney et al., 1980)

Method	The frequency of the method in the literature	Measured State	An indicator of the state	References
Head movements	1	Drowsiness	Head movements dispersion	(Popieul et al., 2003)
Oximetry	1	Fatigue	K-means classifier	(Sharma & Bundele, 2015)
Blood glucose	1	Mental workload	A decline of the blood glucose	(Fairclough & Houston, 2004)

14. APPENDIX 3- TORS QUESTIONNAIRE

Please choose the sentence that best describes how ready do you feel to take-over the manual driving during next minute.

1. Could not be more ready

2. Fully ready

3. Somewhat ready

4. Not sure if I am ready

5. Somewhat not ready

6. I do not feel ready

7. I do not feel able to drive manually at all

15. APPENDIX 4- EEG BRANDS COMPARISON

Table 15.1: List of the papers using Mobita-32 EEG device in the experiment with some technical properties of the EEG set.

Studies that used Mobita-32	**Studied function**	**Data quality comments**	**Models compared with Mobita**	**Result of comparison**	**Additional comments**
(Askamp & van Putten, 2014)	Usability in at-home EEG recording in epilepsy patients	n/a	Trea, Trackit and Safiro	The only comparison of some technical parameters like size, battery type etc., not data quality	

Studies that used Mobita-32	Studied function	Data quality comments	Models compared with Mobita	Result of comparison	Additional comments
(van Erp, Hogervorst, & van der Werf, 2016)	Recording emotional state of the writer while writing a book	n/a	n/a	n/a	Authors achieved a high accuracy of a state classification
(Pinegger, Wriessnegger, Faller, & Müller-Putz, 2016)	Comparison of 3 mobile EEG devices in a BCI P300 spelling device	Mobita had the lowest noise level, had 93% mean accuracy of classification in comparison to 96% in gGammasys and 77% in gSahara, moderate interparticipant variance in comparison with low for gGammasys and high for gSahara	gGammasys and gSahara		Mobita was perceived as most satisfactory by participants and researchers
(Bateson, Baseler, Paulson, Ahmed, & Asghar, 2017)	Classification of the level of mobility of various mobile EEG devices	n/a	20 different mobile EEG models	Mobita was classified as middle device mobility, high participant mobility, and high system specification	

Table 15.2: List of the papers using Enobio-20 EEG device in the experiment with some technical properties of the EEG set.

Studies that used Enobio 20	Studied function	Data quality comments	Models compared with Enobio	Result of comparison	Additional comments
(Mohamed et al., 2018)	Different models of EEG signal analysis in left/right hand/foot EEG recognition	n/a	n/a	n/a	Up to 97.62% accuracy of classification
(Ingle & Awale, 2018)	Assessing effect of Vipasana meditation on brain	n/a	n/a	n/a	Up to 85% accuracy of classification

Studies that used Enobio 20	Studied function	Data quality comments	Models compared with Enobio	Result of comparison	Additional comments
(Angulo-Sherman, Rodríguez-Ugarte, Iáñez, Ortiz, & Azorín, 2017)	Motor imagery with and without TDCS stimulation	n/a	n/a	n/a	Accurate classification of central areas activity
(Rodríguez-Ugarte, Iáñez, Ortíz, & Azorín, 2017)	EEG based BCI	n/a	n/a	n/a	55.1 % of classification
(Ratti, Waninger, Berka, Ruffini, & Verma, 2017)	Comparison of 4 EEG models	All models were assessed as acceptable from the point of view of data quality. B-Alert and Enobio were viewed as better	B-Alert, Muse and Mindwave	B-Alert, Enobio and Mindwave have comparable Fp1 power spectra; Muse has higher test-retest variation, Muse and Mindwave are more susceptible to movement artefacts, Muse and Mindwave are quicker to set-up	
(Sharma, Jain, & Pal, 2017)	Manipulation of robotic arm with EEG extracted artifacts	n/a	n/a	n/a	High classification accuracy
(Biswas et al., 2016)	Investigation of EEG reactions to the changing environment in the working place: temperature, humidity etc.	n/a	n/a	n/a	
(Placidi, Petracca, Spezialetti, & Iacoviello, 2016)	Assessment of BCI technique and interface based on olfactory imagery	n/a	n/a	n/a	Up to 95% accuracy
(Rodríguez-Ugarte et al., 2016)	Detection of pedalling intention with EEG	n/a	n/a	n/a	Max 72.2% of accuracy
(Pistoia et al., 2015)	Detection of olfactory imagery in minimally conscious state	n/a	n/a	n/a	~70% of imagery classification

Studies that used Enobio 20	Studied function	Data quality comments	Models compared with Enobio	Result of comparison	Additional comments
(Awais, Badruddin, & Drieberg, 2014)	Observing EEG frequency changes during monotonous driving	n/a	n/a	n/a	Significant increase of alpha and theta after monotonous driving
(Abbate, Avvenuti, & Light, 2012)	The idea of the minimally invasive system of sensors for early detection of brain problems (with Enobio for EEG)	n/a	n/a	n/a	Researchers assess the system as a reliable tool used when a patient is active

Table 15.3: List of the papers using EEGO Sport EEG device in the experiment with some technical properties of the EEG set.

Studies that used EEGO Sport	Studied function	Data quality comments	Models compared with EEGO Sport	Result of comparison	Additional comments
(Chen et al., 2018)	Assessment of the model of FCMC	n/a	n/a	n/a	
(Stone, Tamburro, Fiedler, Haueisen, & Comani, 2018)	Testing a new algorithm for physiological noise removal	n/a	ECI electrodes system with dry electrodes	n/a	90% of noise successfully removed
(Hall, Mattingley, & Dux, 2018)	EEG activity related to expected and unexpected stimuli	n/a	n/a	n/a	95.8% accuracy for EEG and ERP
(Bigliassi, Karageorghis, Hoy, & Layne, 2018)	Investigation of EEG during positive mental states elicited by exercises and music	n/a	n/a	n/a	Successful oscillations recording during running
(Yue Dong, Raif, Determan, & Gai, 2017)	EEG based detection of auditory attention	n/a	n/a	n/a	Accuracy of classification 70%-90%

Studies that used EEGO Sport	Studied function	Data quality comments	Models compared with EEGO Sport	Result of comparison	Additional comments
(Fehér, Nakataki, & Morishima, 2017)	EEG recording in DLFPC during TMS and tACS	n/a	n/a	n/a	
(Griffiths, Mazaheri, Debener, & Hanslmayr, 2016)	Studying EEG mechanisms of memory formation during the walk around the campus	P300 recorded in the lab was comparable to P300 recorded outdoors	n/a	n/a	Participants were not walking during EPOCS of EEG recording

16. APPENDIX 5- UNIPHYSIOLOGICAL REGRESSION ANALYSIS FOR TAKE-OVER MONITORING, TABLES WITH ALL TESTED MODELS

Table 16.1: 1st linear model predicting change in driving performance based on change in questionnaires' results: Number of observations: 79, Error degrees of freedom: 67, Root Mean Squared Error: 169, R-squared: .15, Adjusted R-Squared: .01, F-statistic vs. constant model: 1.04, p value: NS

	Estimate	SE	t-Stat	P Value
Intercept	-135.64	59.57	-2.28	.03
Circadian Phase	-10.17	41.58	-0.24	.81
Chronotype	-0.29	25.59	-0.01	.99
Change in KSS	-12.90	15.73	-0.82	.42

	Estimate	SE	t-Stat	P Value
Chane in TORS	11.46	18.37	0.62	.54
Change in Fatigue	-37.88	29.97	-1.26	.21
Change in NASA-TLX MD	0.85	0.74	1.15	.26
Change in NASA-TLX PD	0.02	1.33	0.02	.99
Change in NASA-TLX TD	-0.17	1.24	-0.13	.89
Change in NASA-TLX Per	1.21	1.06	1.14	.26
Change on NASA-TLX Eff	0.24	0.79	0.31	.76
Change in NASA-TLX Frust	-0.53	0.71	-0.75	.46
Change in NASA-TLX Overall	0	0	NaN	NaN

Table 16.2: 2nd linear model predicting change in driving performance based on change in questionnaires' results: Number of observations: 81, Error degrees of freedom: 76, Root Mean Squared Error: 163, R-squared: 0.12, Adjusted R-Squared: .07, F-statistic vs. constant model: 2.52, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	-125.10	40.77	-3.07	.00
Change in KSS	-11.91	13.90	-0.86	.39
Change in Fatigue	-26.61	25.04	-1.06	.29
Change in NASA-TLX MD	1.14	0.62	1.84	.07
Change in NASA-TLX Eff	1.19	0.92	1.29	.20

Table 16.3: 2nd linear model predicting change in driving performance based on change in questionnaires' results: Number of observations: 84, Error degrees of freedom: 81, Root Mean Squared Error: 164, R-squared: .06, Adjusted R-Squared: .03, F-statistic vs. constant model: 2.47, p value : NS

	Estimate	SE	t-Stat	P Value
Intercept	-51.20	23.13	-2.21	.03
Change in NASA-TLX MD	1.22	0.61	2.01	.05
Change in NASA-TLX Eff	1.07	0.93	1.16	.25

Table 16.4: 1st linear model predicting change in driving performance based on change in ECG measurements: Number of observations: 96, Error degrees of freedom: 81, Root Mean Squared Error: 159, R-squared: .17, Adjusted R-Squared: .02, F-statistic vs. constant model: 1.17, p value : NS

	Estimate	SE	t-Stat	P Value
Intercept	-52.29	-55.12	0.95	.35
Circadian Phase	-30.99	33.94	-0.91	.36
Session	3.08	35.13	0.09	.93
Change in HR	2.08	4.05	0.51	.61
Change in HRVvl	0.00	0.02	0.22	.83
Change in HRVl	-0.00	0.01	-0.18	.86
Change in HRVh	0.00	0.01	0.52	.60
Change in HRVvh	0.01	0.01	2.36	.02
Change in Sympathetic	-171320000	283150000	-0.61	.55
Vagal	-171320000	283150000	-0.61	.55
Change in Sympathetic to Vagal Ratio	10.84	6.78	1.60	.11
Change in RSA	35.83	29.80	1.20	.23
Change in RMSSD	-185.99	295.08	-0.63	.53
Change in SDSD	186	295.09	0.63	.53
Change in pNN50%	-1.33	1.74	-0.77	.45

Table 16.5: 2nd linear model predicting change in driving performance based on change in ECG measurements: Number of observations: 96, Error degrees of freedom: 91, Root Mean Squared Error: 153, R-squared: .14, Adjusted R-Squared: .10, F-statistic vs. constant model: 3.67, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	-36.52	22.55	-1.62	.11
Circadian Phase	-31.213	31.53	-0.99	.33
Change in HRVvh	0.01	0.01	2.57	.01
Change in Sympathetic to Vagal Ratio	7.87	5.27	1.49	.14
Change in RSA	35.29	16.94	2.08	.04

Table 16.6: 3rd linear model predicting change in driving performance based on change in ECG measurements: Number of observations: 96, Error degrees of freedom: 93, Root Mean Squared Error: 154, R-squared: .11, Adjusted R-Squared: .09, F-statistic vs. constant model: 5.68, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	-57.47	16.35	-3.51	.00
Change in HRVvh	0.01	0.01	2.46	.02
Change in RSA	34.37	16.81	2.05	.04

Table 16.7: 1st linear model predicting change in driving performance based on change in acoustic voice properties: Number of observations: 96, Error degrees of freedom: 76, Root Mean Squared Error: 16, R-squared: .22, Adjusted R-Squared: .02, F-statistic vs. constant model: 1.11, p value : NS

	Estimate	SE	t-Stat	P Value
Intercept	-55.117	57.32	-0.96	.34
Session	-3.25	34.58	-0.09	.93
Circadian Phase	-19.06	37.462	-0.51	.61
Change in Mean Pitch	-1.92	1.80	-1.09	.28
Change in Max Pitch	0.15	0.74	0.20	.85
Change in Min Pitch	0	0	NaN	NaN
Change in Frequency Range	-0.61	0.75	-0.81	.42
Change in Pitch SD	4.22	1.35	3.12	.00
Change in Number of Pulses	-4.91	3.91	-1.25	.21
Change in Number of Periods	5.77	3.94	1.47	.15
Change in Fraction of Locally Unvoiced Frames	2.58	2.95	0.88	.38
Change in Number of Voice Breaks	-16.71	12.18	-1.37	.17
Change in Degree of Voice Breaks	-1.04	2.49	-0.42	.68
Change in Jitter	-39.68	29.92	-1.33	.19
Change in Shimmer	-7.99	10.25	-0.78	.44
Change in Mean Autocorrelation	-2443.6	6325.4	-0.39	.70
Change in Mean noise to harmonics ratio	-950.56	2655.3	-0.36	.72
Change in Mean harmonics to noise ratio	-2.58	42.41	-0.06	.95

	Estimate	SE	t-Stat	P Value
Change in Mean Intensity	-4.27	5.32	-0.80	.43
Change in Max Intensity	0.25	0.44	0.55	.58
Change in Min Intensity	6.85	4.22	1.63	.11

Table 16.8: 2nd linear model predicting change in driving performance based on change in acoustic voice properties: Number of observations: 97, Error degrees of freedom: 92
Root Mean Squared Error: 159, R-squared: .06, Adjusted R-Squared: .02, F-statistic vs. constant model: 1.47, p value : NS

	Estimate	SE	t-Stat	P Value
Intercept	-64.22	17.56	-3.66	.00
Change in Pitch SD	0.33	0.63	0.53	.60
Change in Number of Periods	0.18	0.36	0.51	.61
Change in Number of Breaks	-8.16	4.40	-1.86	.07
Change in Min Intensity	3.62	3.65	0.99	.32

Table 16.9: 3rd linear model predicting change in driving performance based on change in acoustic voice properties: Number of observations: 97, Error degrees of freedom: 94, Root Mean Squared Error: 158, R-squared: .05, Adjusted R-Squared: .03, F-statistic vs. constant model: 2.67, p value : NS

	Estimate	SE	t-Stat	P Value
Intercept	-61.15	16.39	-3.73	.00
Change in Number of Breaks	-7.31	3.39	-2.16	.03
Change in Min Intensity	3.66	3.60	1.02	.31

Table 16.10: 1st linear model predicting change in driving performance based on change in EOG variables: Number of observations: 21, Error degrees of freedom: 13, Root Mean Squared Error: 182, R-squared: .24, Adjusted R-Squared: -.16, F-statistic vs. constant model: 0.6, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	95.58	135.88	0.70	.49
Circadian Phase	-95.76	94.22	-1.07	.33
Session	-103.54	102.82	-1.01	.33
Change in Blinking Rate	4.66	8.77	0.53	.60
Change in Mean Blink Duration	432.9	832.81	0.52	.61
Change in PERCLOS	-24.67	23.04	-1.07	.30
Change in Rate of Horizontal eye-movements	-1.20	9.10	-0.13	.90
Change in Mean Duration of	37.86	189.18	0.20	.85

	Estimate	SE	t-Stat	P Value
Horizontal Eye-movements				

Table 16.11: 2nd linear model predicting change in driving performance based on change in EOG variables: Number of observations: 45, Error degrees of freedom: 41, Root Mean Squared Error: 153, R-squared: .13, Adjusted R-Squared: .07, F-statistic vs. constant model: 2.02, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	91.21	76.14	1.20	.24
Session	-76.87	45.91	-1.67	.10
Circadian Phase	-72.82	47.20	- 1.54	.13
Change in PERCLOS	-9.66	7.43	-1.30	.20

Table 16.12: 1st linear model predicting change in driving performance based on change in EDA variables: Number of observations: Number of observations: 87, Error degrees of freedom: 82
Root Mean Squared Error: 160, R-squared: .11, Adjusted R-Squared: .06, F-statistic vs. constant model: 2.42, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	-37.71	54.33	-0.69	.49
Session	-5.71	34.97	-0.16	.87
Circadian Phase	-22.94	34.74	-0.66	.51
Change in SCL Frequency	101660	37	2.68	.01
Change in SCL Mean	9.84	5.15	1.91	.06

Table 16.13: 2nd linear model predicting change in driving performance based on change in EDA variables: Number of observations: 87, Error degrees of freedom: 84, Root Mean Squared Error: 159, R-squared: .1, Adjusted R-Squared: .08, F-statistic vs. constant model: 4.68, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	-56.75	18.39	-3.09	.00
Change in SCL Frequency	103560	37555	2.76	.01
Change in SCL Mean	9.94	5.05	1.97	.05

Table 16.14: 1st linear model predicting change in driving performance based on change in EMG variables: Number of observations: Number of observations: 42, Error degrees of freedom: 31
Root Mean Squared Error: 157, R-squared: .16, Adjusted R-Squared: -.12, F-statistic vs. constant model: 0.58, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	79.53	89.71	0.89	.38
Session	-66.05	52.09	-1.27	.22
Circadian Phase	-29.68	51.75	-0.57	.57
Change in Median Frequency of Frontalis	0.37	0.92	0.40	.69

	Estimate	SE	t-Stat	P Value
Change in Mean Frequency of Frontalis	0.22	0.56	0.40	.69
Change in Peak Frequency of Frontalis	-1.91	2.17	-0.88	.39
Change in Mean Power of Frontalis	0	0	NaN	NaN
Change in Total Power of Frontalis	-29178000	42427000	-0.69	0.50
Change in Median Frequency of Corrugator Supercilii	-0.52	0.83	-0.62	0.54
Change in Mean Frequency of Corrugator Supercilii	-0.05	0.60	-0.08	0.94
Change in Peak Frequency of Corrugator Supercilii	2.56	2.65	0.97	0.34
Change in Mean Power of Corrugator Supercilii	0	0	NaN	NaN
Change in Total Power of Corrugator Supercilii	-44181000	37595000	-1.18	0.25

Table 16.15: 2nd linear model predicting change in driving performance based on change in EMG variables: Number of observations: 42, Error degrees of freedom: 38, Root Mean Squared Error: 149, R-squared: .06, Adjusted R-Squared: -.01, F-statistic vs. constant model: 0.85, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	53.27	76.56	0.70	.49
Session	-58.45	46.81	-1.25	.22
Change in Peak Frequency in Frontalis	0.04	0.72	0.05	.96
Change in Total Power in Corrugator Supercilii	-32531000	29778000	-1.09	.28

Table 16.16: 1st linear model predicting change in driving performance based on change in respiration variables: Number of observations: 95, Error degrees of freedom: 90, Root Mean Squared Error: 162, R-squared: .04, Adjusted R-Squared: -.00, F-statistic vs. constant model: 0.94, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	-52.56	53.48	-0.98	.33

	Estimate	SE	t-Stat	P Value
Session	-4.07	33.48	-0.12	.90
Circadian Phase	-23.19	33.68	-0.69	.49
Respiration Rate	-11.48	6.86	-1.68	.10
SD of breath	3.94	22.61	0.17	.86

Table 16.17: 1st linear model predicting change in driving performance based on change in salivary hormonal content: Number of observations: 40, Error degrees of freedom: 35, Root Mean Squared Error: 157, R-squared: .03, Adjusted R-Squared: -.08, F-statistic vs. constant model: 0.28, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	-58.07	80.95	-0.72	.48
Session	20.89	52.15	0.40	.69
Circadian Phase	-49.32	51.50	-0.96	.35
Change in Cortisol	-66.81	217.52	-0.31	.76
Change in Alpha-Amylase	0.11	0.28	0.39	.70

Table 16.18: 1st linear model predicting change in driving performance based on change in salivary hormonal content: Number of observations: 94, Error degrees of freedom: 89, Root Mean Squared Error: 163, R-squared: .04, Adjusted R-Squared: -.00, F-statistic vs. constant model: 0.96, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	-59.34	55.20	-1.075	.29
Session	4.05	34.98	0.12	.91
Circadian Phase	-20.34	33.85	-0.60	.55
Change in Sp02	18.74	16.75	1.12	.27
Change in Pulse	-1.82	1.23	-1.48	.14

17. APPENDIX 6- RHO-SPEARMAN CORRELATION TABLES BETWEEN PHYSIOLOGICAL FACTORS AND FACTORS IN MANUAL DRIVING PERFORMANCE FOLLOWING THE MEASUREMENT

Table 17.1: Bonferroni corrected significant correlations between questionnaires results p<.05 and following driving performance.

	Pedestrian Hit	**Speeding Tickets**	**Traffic Lights violation**	**Centreline Crossings**	**Road excursions**	**SD Lane position**	**SD Steering Wheel Angle**	**SD Vehicle Heading Angle**	**SD Longitudinal Acceleration**	**SD longitudinal speed**	**General Driving Perf.**
KSS rs (200)							.19	.15		.17	
TORS rs (181)		-.15						.15			
Fatigue rs (207)					-.18		.17				
NASA-TLX MD rs (179)	22		.17	.18		.16	.29	.17		.18	.18
NASA-TLX TD rs(179)	.16			.16	.20			.15			
NASA-TLX Eff rs(179)							.29	.26		.18	
NASA-TLX Frust rs(179)							.22	.34			
NASA-TLX Overall rs(179)							.34	.35		19	.22

Table 17.2: Bonferroni corrected significant correlations (p<.05) between ECG variables and driving performance variables.

	Number of collisions	**Road Excursions**	**SD lane position**	**SD of the steering wheel angle**	**SD of the vehicle heading angle**	**SD longitudinal acceleration**	**SD of the longitudinal speed**	**General Driving Performance**
HRV very low frequencies rs(202)				.18	.26		.15	.16
HRV low frequencies rs(202)				.17	.29		.15	.17
HRV high frequencies rs(202)	.19		.16	.14				.17
HRV very high frequencies rs(202)	.16		.16	.17	.21	.17	.16	.21
HRV sympathetic tonus rs(202)	-.18							
HRV vagal tonus rs(202)	.18							
HRV sympathetic vagal tonus ratio rs(202)	-.18							
HRV RSA rs(202)	.19		.16	.14				.17
HRV RMSSD rs(202)		-.18		.17			.21	

	Number of collisions	Road Excursions	SD lane position	SD of the steering wheel angle	SD of the vehicle heading angle	SD longitudinal acceleration	SD of the longitudinal speed	General Driving Performance
HRV SDSD rs(202)		-.18		.17			.21	
HRV PNN50 rs(202)		-.21		.14			.14	

Table 17.3: Bonferroni corrected significant correlations (p<.05) between acoustic voice properties and driving performance, as well as some additional correlations between demografic variables and driving performance.

	Collisions	Pedestrian Hit	Speed Exceedances	Speeding Tickets	Centreline Crossing	Road Excursions	SD of Lane Position	SD of steering wheel angle	SD of Vehicle Heading Angle	SD of longitudinal acceleration	SD of longitudinal speed	Summary of the Driving Performance
Mean Pitch rs(203)			-.27	-.28	-.18		-.19	-.37		-.28	-.3	-.31
Max Pitch rs(203)		.14	-.15					-.19			-.15	-.17
Frequency Range rs(203)								-.16				-.15
Pitch SD rs(203)			-.18	-.17	-.15		-.14	-.24		-.2	-.22	-.24
Number of Pulses rs(203)			-.3	r-.31	-.24		-.17	-.39	-.16	-.29	-.31	-.35
Number of Periods rs(203)			-.28	-.3	-.23		-.17	-.37	-.16	-.28	-.30	-.33
Fraction of Locally Unvoi											-.22	

	Collisions	**Pedestrian Hit**	**Speed Exceedances**	**Speeding Tickets**	**Centreline Crossing**	**Road Excursions**	**SD of Lane Position**	**SD of steering wheel angle**	**SD of Vehicle Heading Angle**	**SD of longitudinal acceleration**	**SD of longitudinal speed**	**Summary of the Driving Performance**
ced Frames rs(203)												
Number of Breaks rs(203)			-.20	-.18	-.17	-.23		-.27		-.24	-.29	-.3
Degree of Breaks rs(203)						-.14						
Jitter rs(203)	.14		.17						.19			.14
Shimmer rs(203)								.18		.14	.16	.04
Mean Harmonics to Noise Ration rs(203)								-.19			-.14	
Mean Intensity rs(203)		.15	.17							.19		
Max Intensity rs(203)										.17		

Table 17.4: Bonferroni corrected significant correlations (p<.05) between EOG variables and factors of driving performance.

	Collisions	**Centreline Crossing**	**SD of longitudinal acceleration**
Blinking Rate rs(92)	.37	.35	.31

	Collisions	**Centreline Crossing**	**SD of longitudinal acceleration**
Mean Blink Duration rs(92)			.03 -.29
PERCLOS rs(92)		.34	

Table 17.5: Bonferroni corrected significant correlations (p<.05) between respiration factors and factors of driving performance.

	Stop Signs Violations	**Traffic Lights Violations**	**SD of lane position**
Breathing Rate rs(200)		.25	.15
Standard Deviation of Breath rs(200)	.16		-.18

18. APPENDIX 7- UNIPHYSIOLOGICAL REGRESSION MODELS PREDICTING ATTENTION DURING AUTOMATION WITH PHYSIOLOGICAL RECORDING COLLECTED BEFORE THE AUTOMATED MODE, TABLES WITH ALL TESTED MODELS

Table 18.1: 1st linear model predicting attention during automated mode based on the questionnaires collected just before automation: Number of observations: 80, Error degrees of freedom: 68, Root Mean Squared Error: 0.94, R-squared: .102, Adjusted R-Squared: -.04, F-statistic vs. constant model: 0.7, p value : NS

	Estimate	SE	t-Stat	P Value
Intercept	4.51	0.59	7.61	.00
Circadian Phase	0.38	0.25	1.54	.13
Session	0.03	0.24	0.14	.89
KSS	0.05	0.12	0.42	.68
TORS	0.10	0.14	0.73	.47
Fatigue	-0.33	0.19	-1.73	.09
NASA-TLX MD	0.01	0.01	0.92	.36
NASA-TLX PD	-0.01	0.01	-0.60	.55
NASA-TLX TD	-0.00	0.01	-0.49	.63
NASA-TLX Per	-0.00	0.01	-0.47	.64
NASA-TLX Eff	0.00	0.01	0.50	.62
NASA-TLX Frust	0.00	0.00	0.51	.61
NASA-TLX Overall	0	0	NaN	NaN

Table 18.2: 2nd linear model predicting attention during automated mode based on the questionnaires collected just before automation: Number of observations: 82, Error degrees of freedom: 78, Root Mean Squared Error: 0.91, R-squared: .07, Adjusted R-Squared: .03, F-statistic vs. constant model: 1.85, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	4.61	0.28	16.67	.00
Circadian Phase	0.38	0.22	1.72	.09
TORS	0.10	0.13	0.76	.45
Fatigue	-0.22	0.10	-2.09	.04

Table 18.3: 3rd linear model predicting attention during automated mode based on the questionnaires collected just before automation: Number of observations: 91, Error degrees of freedom: 88, Root Mean Squared Error: 0.885, R-squared: 0.07, Adjusted R-Squared: .05, F-statistic vs. constant model: 3.46, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	4.79	0.23	21.33	.00
Circadian Phase	0.38	0.20	1.89	.06

	Estimate	SE	t-Stat	P Value
Fatigue	-0.19	0.08	-2.44	.02

Table 18.4: 4th linear model predicting attention during automated mode based on the questionnaires collected just before automation: Number of observations: 91, Error degrees of freedom: 89, Root Mean Squared Error: 0.90, R-squared: .04, Adjusted R-Squared: .03, F-statistic vs. constant model: 3.27, p Value:NS

	Estimate	SE	t-Stat	P Value
Intercept	4.80	0.23	21.11	.00
Fatigue	-0.13	0.07	-1.81	.07

Table 18.5: 1st linear model predicting attention during automated mode based on the ECG collected in the resting state just before automation: Number of observations: 93, Error degrees of freedom: 77, Root Mean Squared Error: 0.92, R-squared: .12, Adjusted R-Squared: -.05, F-statistic vs. constant model: 0.72, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	-2298400	2822900	-0.81	.42
Circadian Phase	0.05	0.21	0.26	.80
HR	0.02	0.02	1.41	.16
HRVvl	-0.00	0.00	-0.41	.69
HRVl	0.00	0.00	-0.19	.85
HRVh	-0.00	0.00	-0.60	.55
HRVvh	0.00	0.00	1.75	.09
HRV sympathetic tonus	2298400	2822900	0.81	.42
HRV vagal tonus	2298400	2822900	0.81	.42
HRV sympathetic/vagal tonus ration	0.01	0.09	0.09	.93
RSA	-0.28	0.22	-1.27	.21
RMSSD	-0.35	0.38	-0.93	.36
SDSD	0.35	0.38	0.93	.36
pNN50%	0.02	0.01	1.81	.07
Chronotype	0.07	0.12	0.60	.55
Session	0.04	0.20	0.20	.84

Table 18.6: 2nd linear model predicting attention during automated mode based on the ECG collected in the resting state just before automation: Number of observations: 93, Error degrees of freedom: 88, Root Mean Squared Error: 0.89, R-squared: .07, Adjusted R-Squared: .02, F-statistic vs. constant model: 1.53, p Value :NS

	Estimate	SE	t-Stat	P Value
Intercept	4.59	1.08	4.26	.00
HR	0.01	0.01	1.19	.24
HRVvh	0.00	0.00	1.39	.17

	Estimate	SE	t-Stat	P Value
RSA	-0.21	0.12	-1.85	.07
pNN50%	0.01	0.01	1.79	.08

Table 18.7: 1st linear model predicting attention during automated mode based on the acoustic voice properties from the voice recording collected just before the automated mode: Number of observations: 93, Error degrees of freedom: 72, Root Mean Squared Error: 0.94, R-squared: .15, Adjusted R-Squared: -.09, F-statistic vs. constant model: 0.61, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	45.44	44.50	1.02	.31
Chronotype	0.11	0.15	0.78	.44
Session	0.03	0.21	0.13	.90
Circadian Phase	0.14	0.21	0.66	.51
Mean Pitch	-0.00	0.01	-0.31	.76
Max Pitch	-0.01	0.01	-0.97	.34
Min Pitch	0	0	NaN	NaN
Frequency Range	0.01	0.01	1.18	.24
Pitch SD	-0.01	0.01	-0.84	.40
Number of Pulses	0.00	0.02	0.43	.67
Number of Periods	-0.01	0.03	-0.28	.78
Fraction of Locally Unvoiced Frames	-0.00	0.02	-0.11	.91
Number of Voice Breaks	-0.09	0.07	-1.30	.20
Degree of Voice Breaks	0.01	0.02	0.24	.81
Jitter	-0.19	0.23	-0.85	.40
Shimmer	0.05	0.07	0.77	.45
Mean Autocorrelation	-46.27	48.74	-0.95	.35
Mean noise to harmonics ratio	-15.51	22.02	-0.70	.48
Mean harmonics to noise ratio	0.26	0.27	0.96	.34
Mean Intensity	0.09	0.08	1.06	.29
Max Intensity	-0.06	0.08	-0.68	.50
Min Intensity	-0.05	0.05	-1.16	.25

Table 18.8: 2nd linear model predicting attention during automated mode based on the acoustic voice properties from the voice recording collected just before the automated mode: Number of observations: 93, Error degrees of freedom: 88, Root Mean Squared Error: 0.90, R-squared: .06, Adjusted R-Squared: .02, F-statistic vs. constant model: 1.36, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	4.68	1.20	3.89	.00
Number of Breaks	0.00	0.00	0.80	.42
Frequency Range	-0.08	0.05	-1.68	.10
Mean Intensity	0.02	0.02	1.26	.21
Min Intensity	-0.04	0.04	-0.93	.35

Table 18.9: 1st linear model predicting attention during automated mode based on the ocular behaviours measured with EOG recording collected during the resting state just before the automated mode: Number of observations: 32, Error degrees of freedom: 25, Root Mean Squared Error: 0.991, R-squared: .39, Adjusted R-Squared: .24, F-statistic vs. constant model: 2.63, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	5.94	0.76	7.87	.00
Circadian Phase	-0.25	0.42	-0.58	.57
Blink Rate	-0.03	0.04	-0.73	.47
Mean Blink Duration	-4.29	3.34	-1.29	.21
PERCLOS	-0.10	0.15	-0.65	.52
Rate of Horizontal Eye-Movements	0.02	0.03	0.49	.63
Mean Duration of Horizontal Eye-Movements	0.68	0.63	1.07	.29

Table 18.10: 2nd linear model predicting attention during automated mode based on the ocular behaviours measured with EOG recording collected during the resting state just before the automated mode: Number of observations: 32, Error degrees of freedom: 29, Root Mean Squared Error: 1.11, R-squared: .11, Adjusted R-Squared: .05, F-statistic vs. constant model: 1.77, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	4.74	0.47	10.06	.00
Mean Blink Duration	-3.79	2.22	-1.71	.10
Mean Duration of Horizontal Eye-Movements	0.76	0.68	1.11	.28

Table 18.11: 3rd linear model predicting attention during automated mode based on the ocular behaviours measured with EOG recording collected during the resting state just before the automated mode: Number of observations: 45, Error degrees of freedom: 41, Root Mean Squared Error: 0.92, R-squared: .25, Adjusted R-Squared: .20, F-statistic vs. constant model: 4.58, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	4.93	0.22	22.25	.00
Blink Rate	0.00	0.02	0.02	.99
PERCLOS	-0.16	0.08	-1.99	.05
Rate of Horizontal Eye-Movements	0.02	0.03	0.49	.63

Table 18.12: 4th linear model predicting attention during automated mode based on the ocular behaviours measured with EOG recording collected during the resting state just before the automated mode: Number of observations: 45, Error degrees of freedom: 43, Root Mean Squared Error: 0.90, R-squared: .25, Adjusted R-Squared: .23, F-statistic vs. constant model: 14.1, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	4.91	0.20	24.15	.00
PERCLOS	-0.14	0.04	-3.75	.00

Table 18.13: 1st linear model predicting attention during automated mode based on the EDA variables collected during the resting state just before the automated mode: Number of observations: 85, Error degrees of freedom: 81, Root Mean Squared Error: 0.93, R-squared: .03, Adjusted R-Squared: -.01, F-statistic vs. constant model: 0.79, p value<.50

	Estimate	SE	t-Stat	P Value
Intercept	2.78	2.52	1.10	.27
Circadian Phase	0.20	0.21	0.94	.35
SCL Frequency	172.1	320.69	0.54	.59
SCL Mean	0.02	0.02	0.99	.33

Table 18.14: 1st linear model investigating association between attention during automated mode and EDA variables collected during the whole automated mode: Number of observations: 85, Error degrees of freedom: 80, Root Mean Squared Error: 0.93, R-squared: .06, Adjusted R-Squared: .01, F-statistic vs. constant model: 1.16, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	3.92	0.37	10.65	.00
Chronotype	0.10	0.12	0.82	.41
Session	0.04	0.20	0.22	.83
Circadian Phase	0.20	0.20	1.00	.32
SCL Frequency	0	0	NaN	NaN
SCL Mean	0.03	0.02	1.59	.12

Table 18.15:2nd linear model investigating association between attention during automated mode and EDA variables collected during the whole automated mode: Number of observations: 85, Error degrees of freedom: 83, Root Mean Squared Error: 0.92, R-squared: .03, Adjusted R-Squared: .02, F-statistic vs. constant model: 2.83, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	4.15	0.18	23.30	.00
SCL Frequency	0	0	NaN	NaN
SCL Mean	0.03	0.02	1.68	.10

Table 18.16: 1st linear model predicting attention during automated mode based on the EMG variables collected during the resting state just before the automated mode: Number of observations: 43, Error degrees of freedom: 33, Root Mean Squared Error: 0.91, R-squared: .20, Adjusted R-Squared: -.01, F-statistic vs. constant model: 0.94, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	5.09	1.21	4.21	.00
Circadian Phase	0.38	0.30	1.24	.22
Median Frequency of Frontalis	0.00	0.01	0.25	.80
Mean Frequency of Frontalis	0.00	0.00	0.13	.90
Peak Frequency of Frontalis	-0.01	0.01	-1.20	.24

	Estimate	SE	t-Stat	P Value
Mean Power of Frontalis	0	0	NaN	NaN
Total Power of Frontalis	-310300	447120	-0.69	.50
Median Frequency of Corrugator Supercilii	-0.00	0.00	-0.20	.84
Mean Frequency of Corrugator Supercilii	-0.00	0.00	-0.37	.71
Peak Frequency of Corrugator Supercilii	0.01	0.02	0.22	.83
Mean Power of Corrugator Supercilii	0	0	NaN	NaN
Total Power of Corrugator Supercilii	-449460	353140	-1.27	.21

Table 18.17: 2nd linear model predicting attention during automated mode based on the EMG variables collected during the resting state just before the automated mode: Number of observations: 43, Error degrees of freedom: 39, Root Mean Squared Error: 0.87, R-squared: .16, Adjusted R-Squared: .09, F-statistic vs. constant model: 2.41, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	5.03	0.54	9.31	.00
Circadian Phase	0.33	0.27	1.21	.23
Peak Frequency of Frontalis	-0.01	0.01	-1.25	.22
Total Power of Corrugator Supercilii	-496190	225500	-2.20	.03

Table 18.18: 3rd linear model predicting attention during automated mode based on the EMG variables collected during the resting state just before the automated mode: Number of observations: 43, Error degrees of freedom: 41, Root Mean Squared Error: 0.88, R-squared: .09, Adjusted R-Squared: .07, F-statistic vs. constant model: 3.91, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	4.56	0.15	31.22	.00
Total Power of Corrugator Supercilii	-440910	222960	-1.98	.06

Table 18.19: 1st linear model predicting attention during automated mode based on the respiration based variables collected during the resting state just before the automated mode: Number of observations: 93, Error degrees of freedom: 89, Root Mean Squared Error: 0.91, R-squared: .02, Adjusted R-Squared: -.02, F-statistic vs. constant model: 0.44, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	4.18	0.50	8.30	.00
Circadian Phase	0.18	0.19	0.94	.35
Respiration Rate	0.01	0.03	0.26	.80

	Estimate	SE	t-Stat	P Value
SD of Breath	0.04	0.06	0.60	.55

Table 18.20: 1st linear model predicting attention during automated mode based on the oximetry based variables collected during the resting state just before the automated mode: Number of observations: 90, Error degrees of freedom: 84, Root Mean Squared Error: 0.91, R-squared: .04, Adjusted R-Squared: -.02, F-statistic vs. constant model: 0.73, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	3.99	7.14	0.56	.58
Chronotype	0.10	0.12	0.83	.41
Session	0.02	0.19	0.20	.92
Circadian Phase	0.15	0.19	0.75	.46
Sp02	-0.01	0.07	-0.07	.95
Pulse	0.01	0.01	1.21	.23

19. APPENDIX 8- UNIPHYSIOLOGICAL REGRESSION MODELS PREDICTING ATTENTION DURING AUTOMATION WITH PHYSIOLOGICAL RECORDING COLLECTED DURING THE WHOLE AUTOMATED MODE, TABLES WITH ALL TESTED MODELS

Table 19.1: 1st linear model predicting attention during automated mode based on the ECG during the whole period of automation: Number of observations: 93, Error degrees of freedom: 78, Root Mean Squared Error: 0.91, R-squared: .14, Adjusted R-Squared: -.01, F-statistic vs. constant model: 0.91, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	0	0	NaN	NaN
Circadian Phase	0.12	0.20	0.62	.54
HR	0.01	0.02	0.60	.55
HRVvl	-0.00	0.00	-1.55	.13
HRVl	0.00	0.00	0.22	.83
HRVh	0.00	0.00	0.08	.94
HRVvh	0.00	0.00	1.34	.19
HRV sympathetic tonus	5.12	1.89	2.70	.01
HRV vagal tonus	3.51	2.09	1.68	.20
HRV sympathetic/vagal tonus ration	-0.18	0.14	-1.28	.20
RSA	-0.08	0.28	-0.28	.78
RMSSD	-255.98	663.88	-0.39	.70
SDSD	255.98	663.88	0.39	.70
pNN50%	0.01	0.01	0.59	.56
Chronotype	0.11	0.13	0.85	.40
Session	0.07	0.20	0.34	.73

Table 19.2: 2nd linear model predicting attention during automated mode based on the ECG during the whole period of automation: Number of observations: 93, Error degrees of freedom: 88, Root Mean Squared Error: 0.90, R-squared: .06, Adjusted R-Squared: .01, F-statistic vs. constant model: 1.3, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	5.53	0.64	8.65	.00
HRVvl	-0.00	0.00	-0.76	.45
HRVvh	0.00	0.00	1.41	.16
HRV sympathetic tonus	-0.21	0.12	-1.73	.09
HRV vagal tonus	0.01	0.01	1.47	.15

Table 19.3: 1st linear model investigating association between attention during automated mode and ocular behaviours measured with EOG recording collected during the whole automated mode: Number of observations: 45, Error degrees of freedom: 38, Root Mean Squared Error: 0.87, R-squared: .37, Adjusted R-Squared: .27, F-statistic vs. constant model: 3.7, p value <.05

	Estimate	**SE**	**t-Stat**	**P Value**
Intercept	7.34	1.03	7.14	.00
Circadian Phase	0.40	0.28	1.45	.16
Blink Rate	-0.12	0.05	-2.76	.01
Mean Blink Duration	-8.84	3.20	-2.78	.01
PERCLOS	0.27	0.13	2.18	.04
Rate of Horizontal Eye-Movements	0.05	0.03	2.06	.05
Mean Duration of Horizontal Eye-Movements	-4.93	1.75	-2.81	.01

Table 19.4: 2nd linear model investigating association between attention during automated mode and ocular behaviours measured with EOG recording collected during the whole automated mode: Number of observations: 45, Error degrees of freedom: 39, Root Mean Squared Error: 0.89, R-squared: .33, Adjusted R-Squared: .25, F-statistic vs. constant model: 3.9, p value <.05

	Estimate	**SE**	**t-Stat**	**P Value**
Intercept	7.54	1.03	7.29	.00
Blink Rate	-0.12	0.05	-2.64	.01
Mean Blink Duration	-8.71	3.24	-2.69	.01
PERCLOS	0.28	0.13	2.17	.04
Rate of Horizontal Eye-Movements	0.05	0.03	1.72	.09
Mean Duration of Horizontal Eye-Movements	-5.02	1.78	-2.83	.01

Table 19.5: 3rd linear model investigating association between attention during automated mode and ocular behaviours measured with EOG recording collected during the whole automated mode: Number of, observations: 45, Error degrees of freedom: 40, Root Mean Squared Error: 0.91, R-squared: .28, Adjusted R-Squared: .21, F-statistic vs. constant model: 3.94, p value <.05

	Estimate	**SE**	**t-Stat**	**P Value**
Intercept	7.85	1.04	7.52	.00
Blink Rate	-0.14	0.05	-3.16	.00
Mean Blink Duration	-10.58	3.13	-3.38	.00
PERCLOS	0.39	0.11	3.39	.00
Mean Duration of Horizontal Eye-Movements	-4.18	1.75	-2.39	.03

Table 19.6: 1st linear model predicting attention during automated mode based on the EMG variables collected during the resting state just before the automated mode: Number of observations: 41, Error degrees

of freedom: 31, Root Mean Squared Error: 0.89, R-squared: .28, Adjusted R-Squared: .07, F-statistic vs. constant model: 1.34, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	-0.61	2.26	-0.27	.79
Circadian Phase	0.44	0.29	1.50	.15
Median Frequency of Frontalis	-0.01	0.01	-0.90	.37
Mean Frequency of Frontalis	0.01	0.00	2.26	.03
Peak Frequency of Frontalis	0.04	0.03	1.30	.20
Mean Power of Frontalis	0	0	NaN	NaN
Total Power of Frontalis	221550	581120	0.38	.71
Median Frequency of Corrugator Supercilii	-0.00	0.01	-0.09	.93
Mean Frequency of Corrugator Supercilii	0.00	0.01	0.57	.57
Peak Frequency of Corrugator Supercilii	0.01	0.03	0.31	.76
Mean Power of Corrugator Supercilii	0	0	NaN	NaN
Total Power of Corrugator Supercilii	925700	904750	1.02	.32

Table 19.7: 2nd linear model predicting attention during automated mode based on the EMG variables collected during the resting state just before the automated mode: Number of observations: 42, Error degrees of freedom: 40, Root Mean Squared Error: 0.86, R-squared: .14, Adjusted R-Squared: .12, F-statistic vs. constant model: 6.38, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	3.15	0.52	6.05	.00
Mean Frequency of Frontalis	0.01	0.00	2.53	.02

Table 19.8: 1st linear model predicting attention during automated mode based on the respiration variables collected during the whole automated mode: Number of observations: 93, Error degrees of freedom: 87, Root Mean Squared Error: 0.89, R-squared: .08, Adjusted R-Squared: .03, F-statistic vs. constant model: 1.49, p value :NS

	Estimate	SE	t-Stat	P Value
Intercept	2.70	0.71	3.78	.00
Chronotype	0.09	0.11	0.79	.43
Session	0.01	0.19	0.08	.94
Circadian Phase	0.11	0.19	0.58	.57
Respiration Rate	0.10	0.04	2.38	.02

	Estimate	SE	t-Stat	P Value
SD of Breath	0.04	0.07	0.60	.55

Table 19.9: 2nd linear model predicting attention during automated mode based on the respiration variables collected during the whole automated mode: Number of observations: 93, Error degrees of freedom: 91, Root Mean Squared Error: 0.88, R-squared: .06, Adjusted R-Squared: .05, F-statistic vs. constant model: 6.12, p value <.05

	Estimate	SE	t-Stat	P Value
Intercept	2.89	0.63	4.60	.00
Respiration Rate	0.09	0.04	2.47	.02

20. APPENDIX 9- UNIPHYSIOLOGICAL BINOMIAL REGRESSION MODELS PREDICTING RESULTS OF THE ATTENTION TEST DURING AUTOMATION WITH PHYSIOLOGICAL RECORDING COLLECTED DURING PRECEDING 30-SECOND PERIOD, TABLES WITH ALL TESTED MODELS

Table 20.1: 1st binomial model predicting detection of the red car based on the ECG recorded during 30 seconds before red car appeared in the simulation: Error degrees of freedom: 453, R-squared: .02, Adjusted R-Squared: .00.

	Estimate	SE	t-Stat	P Value
Intercept	6393383.80	4256291.22	1.50	.13
HR	0.03	0.02	1.94	.05
HRVvl	0.00	0.00	0.61	.54
HRVl	0.00	0.00	0.47	.64
HRVh	-0.00	0.00	-0.93	.35
HRVvh	0.00	0.00	1.92	.06
HRV sympathetic tonus	-6393416.29	4256301.50	-1.50	.13
HRV vagal tonus	-6393378.09	4256290.40	-1.50	.13
HRV sympathetic/vagal tonus ration	20.97	14.54	1.44	.15
RSA	-0.70	0.28	-2.47	.01
RMSSD	0.01	0.02	0.27	.79
SDSD	-0.01	0.02	-0.27	.79
pNN50%	0.02	0.01	1.38	.17

Table 20.2: 2nd binomial model predicting detection of the red car based on the ECG recorded during 30 seconds before red car appeared in the simulation: Error degrees of freedom: 462, R-squared: .05, Adjusted R-Squared: .05.

	Estimate	SE	t-Stat	P Value
Intercept	2.86	1.82	1.57	.12
HR	0.03	0.02	1.77	.08
HRVvh	0.00	0.00	1.91	.06
RSA	-0.42	0.19	-2.23	.03

Table 20.3: 3rd binomial model predicting detection of the red car based on the ECG recorded during 30 seconds before red car appeared in the simulation: Error degrees of freedom: 463, R-squared: .05, Adjusted R-Squared: .04.

	Estimate	SE	t-Stat	P Value
Intercept	5.47	1.13	4.82	.00
HRVvh	0.00	0.00	2.21	.03
RSA	-0.56	0.18	-3.18	.00

Table 20.4: 1st binomial model predicting detection of the red car based on the ocular behaviours recorded with EOG during 30 seconds before red car appeared in the simulation: Error degrees of freedom: 137, R-squared: .36, Adjusted R-Squared: .34.

	Estimate	SE	t-Stat	P Value
Intercept	2.45	1.15	2.13	.03
Blink Rate	-0.01	0.05	-0.30	.77
Mean Blink Duration	0.72	3.03	0.24	.81
PERCLOS	-0.09	0.13	-0.64	.52
Rate of Horizontal Eye-Movements	0.06	0.05	1.23	.22
Mean Duration of Horizontal Eye-Movements	1.31	2.13	0.61	.54

Table 20.5: 2nd binomial model predicting detection of the red car based on the ocular behaviours recorded with EOG during 30 seconds before red car appeared in the simulation: Error degrees of freedom: 140, R-squared: .36, Adjusted R-Squared: .35.

	Estimate	SE	t-Stat	P Value
Intercept	2.58	0.59	4.36	.00
PERCLOS	-0.07	0.06	-1.15	.25
Rate of Horizontal Eye-Movements	0.05	0.04	1.18	.24

Table 20.6: 1st binomial model predicting detection of the red car based on the ocular behaviours recorded with EOG during 30 seconds before red car appeared in the simulation: Error degrees of freedom: 423, R-squared: .01, Adjusted R-Squared: .01.

	Estimate	SE	t-Stat	P Value
Intercept	2.24	2.64	0.85	.40
SCL Frequency	-23.99	81.14	-0.30	.77
SCL Mean	0.07	0.03	2.36	.02

Table 20.7: 2nd binomial model predicting detection of the red car based on the ocular behaviours recorded with EOG during 30 seconds before red car appeared in the simulation: Error degrees of freedom: 424, R-squared: .01, Adjusted R-Squared: .01

	Estimate	SE	t-Stat	P Value
Intercept	1.46	0.25	5.78	.00
SCL Mean	0.07	0.03	2.36	.02

Table 20.8: 1st binomial model predicting detection of the red car based on the EMG of corrugator supercilii and frontalis muscles recorded during 30 seconds before red car appeared in the simulation: Error degrees of freedom: 202, R-squared: .74, Adjusted R-Squared: .73. Model was overparametrized.

	Estimate	SE	t-Stat	P Value
Intercept	-3.51	2.04	-1.72	.09
Median Frequency of Frontalis	-0.00	0.01	-0.28	.78
Mean Frequency of Frontalis	0.01	0.01	1.93	.05
Peak Frequency of Frontalis	0.01	0.02	0.46	.64
Mean Power of Frontalis	0	0	NaN	NaN
Total Power of Frontalis	671363.12	1030499.35	0.65	.52
Median Frequency of Corrugator Supercilii	0.01	0.01	0.59	.55
Mean Frequency of Corrugator Supercilii	0.01	0.01	1.08	.28
Peak Frequency of Corrugator Supercilii	0.00	0.02	0.13	.89
Mean Power of Corrugator Supercilii	0	0	NaN	NaN
Total Power of Corrugator Supercilii	6971138.35	3962619.11	1.76	.08

Table 20.9: 2nd binomial model predicting detection of the red car based on the EMG of corrugator supercilii and frontalis muscles recorded during 30 seconds before red car appeared in the simulation: Error degrees of freedom: 208, R-squared: .06, Adjusted R-Squared: .06.

	Estimate	SE	t-Stat	P Value
Intercept	-0.40	0.68	-0.59	.55

	Estimate	SE	t-Stat	P Value
Mean Frequency of Frontalis	0.01	0.00	3.12	.00
Total Power of Corrugator Supercilii	3393865.24	1721661.58	1.97	.05

Glossary of Terms

Circadian rhythm- physiological rhythm related to the day and night cycle.

Electrocardiography- a measure of the inhibitory and excitatory postsynaptic potentials of the cortical nerve cells.

Electroencephalography- a measure of electrical heart activity.

Electromyography- a measure of the electrical muscle's activity.

Electrooculography- a measure of the ocular behaviours through the resting potential of the retina.

Electrodermal activity- a property of human physiology characterised by continuous variations in the electrical properties of the skin.

Heart rate variability- natural phenomenon of the differences in between the time duration in between the consecutive heartbeats.

Mean autocorrelation- in voice context, mean correlation between voice signals separated with a unit of time calculated with the following formula, where x(t) is a speech signal in time, τ is a time lag, and w(t) is a window of time:

$$r_x(\tau) \approx \frac{\int x(t)x(t+\tau)dt}{\int w(t)w(t+\tau)dt}$$

PERCLOS- a measure of ocular behaviour calculated as a ratio between the time that eyelid remains closed or almost closed and the whole given period.

Shimmer- irregularities in voice amplitude.

Speech intensity- squared amplitude of the voice from the beginning of the period until the given point.

Take-over- the transition of the vehicle control from machine to human.

Voice breaks- unvoiced segments of the speech.

VOICE PULSES- THE MEASURE OF GLOTTAL CYCLES DURING THE SPEECH LIST OF REFERENCES

Abe, T., Mollicone, D., Basner, M., & Dinges, D. F. (2014). Sleepiness and safety: Where biology needs technology. *Sleep and Biological Rhythms*, *12*(2), 74–84.

Ackermann, S., & Rasch, B. (2014). Differential Effects of Non-REM and REM Sleep on Memory Consolidation? *Current Neurology and Neuroscience Reports*, *14*(2), 430. https://doi.org/10.1007/s11910-013-0430-8

Aghaei, A. S., Donmez, B., Liu, C. C., He, D., Liu, G., Plataniotis, K. N., Chen, H. Y. W., & Sojoudi, Z. (2016). Smart Driver Monitoring: When Signal Processing Meets Human Factors: In the driver's seat. *IEEE Signal Processing Magazine*, *33*(6), 35–48. https://doi.org/10.1109/MSP.2016.2602379

Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, *36*(7), 623–636.

Ahn, S., Nguyen, T., Jang, H., Kim, J. G., & Jun, S. C. (2016). Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data. *Frontiers in Human Neuroscience*, *10*, 219.

Aidman, E., Chadunow, C., Johnson, K., & Reece, J. (2015). Real-time driver drowsiness feedback improves driver alertness and self-reported driving performance. *Accident Analysis & Prevention*, *81*, 8–13.

Åkerstedt, T., & Kecklund, G. (2001). Age, gender and early morning highway accidents. *Journal of Sleep Research*, *10*(2), 105–110. https://doi.org/10.1046/j.1365-2869.2001.00248.x

Akerstedt, T, Kecklund, G., & Horte, L. G. (2001). Night driving, season, and the risk of highway accidents. *Sleep*, *24*(4), 401–406. https://doi.org/10.1093/sleep/24.4.401

Akerstedt, T., Anund, A., Axelsson, J., & Kecklund, G. (2014). Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. *Journal of Sleep Research*, *23*(3), 242–254.

Akerstedt, T., & Folkard, S. (1995). Validation of the S and C components of the three-process model of alertness regulation. *Sleep*, *18*(1), 1–6.

Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*(1–2), 29–37.

Akerstedt, T., Hallvig, D., Anund, A., Fors, C., Schwarz, J., & Kecklund, G. (2013). Having to stop driving at night because of dangerous sleepiness–awareness, physiology and behaviour. *Journal of Sleep Research*, *22*(4), 380–388. Akerstedt, T., & Kecklund, G. (2001). Age, gender and early morning highway accidents. *Journal of Sleep Research*, *10*(2), 105–110. https://doi.org/10.1046/j.1365-2869.2001.00248.x

Akerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2005). Impaired alertness and performance driving home from the night shift: A driving simulator study. *Journal of Sleep Research*, *14*(1), 17–20.

Allen, R. L., & Mills, D. (2004). *Signal analysis: Time, frequency, scale, and structure*. John Wiley & Sons.

Alpaydin, E. (2014). *Introduction to machine learning*. MIT press.

Amirian, I., Andersen, L. T., Rosenberg, J., & Gögenur, I. (2014). Decreased heart rate variability in surgeons during night shifts. *Canadian Journal of Surgery*, *57*(5), 300.

Andorre, V., & Quéinnec, Y. (1998). Changes in supervisory activity of a continuous process during night and day shifts. *International Journal of Industrial Ergonomics*, *21*(3–4), 179–186.

Andorre-Gruet, V., Queinnec, Y., & Concordet, D. (1998). Three-process model of supervisory activity over 24 hours. *Scandinavian Journal of Work, Environment & Health*, 121–127.

Aranyi, G., Charles, F., & Cavazza, M. (2015). *Anger-based BCI using fNIRS neurofeedback*. 511–521.

Arnedt, J. T., Wilde, G. J. S., Munt, P. W., & MacLean, A. W. (2001). How do prolonged wakefulness and alcohol compare in the decrements they produce on a simulated driving task? *Accident Analysis & Prevention*, *33*(3), 337–344.

Aronson, J. E. (1986). The multiperiod assignment problem: A multicommodity network flow model and specialized branch and bound algorithm. *European Journal of Operational Research*, *23*(3), 367–381.

Artkoski, M., Tommila, J., & Laukkanen, A.-M. (2002). Changes in voice during a day in normal voices without vocal loading. *Logopedics Phoniatrics Vocology*, *27*(3), 118–123.

Averty, P., Athenes, S., Collet, C., & Dittmar, A. (2002). Evaluating a new index of mental workload in real ATC situation using psychophysiological measures. *Digital Avionics Systems Conference, 2002. Proceedings. The 21st*, *2*, 7A4–7A4. internal-pdf://0765043329/Evaluating a new index of mental workload in r.pdf

Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., & Onaral, B. (2012). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage*, *59*(1), 36–47.

Baccouch, R., Zarrouk, N., Chtourou, H., Rebai, H., & Sahli, S. (2015). Time-of-day effects on postural control and attentional capacities in children. *Physiology and Behavior*, *142*, 146–151. https://doi.org/10.1016/j.physbeh.2015.01.029

Bailey, N. R., & Scerbo, M. W. (2007). Automation-induced complacency for monitoring highly reliable systems: The role of task complexity, system experience, and operator trust. *Theoretical Issues in Ergonomics Science*, *8*(4), 321–348.

Balkin, T. J., & Badia, P. (1988). Relationship between sleep inertia and sleepiness: Cumulative effects of four nights of sleep disruption/restriction on performance following abrupt nocturnal awakening. *Biological Psychology*, *27*(3), 245–258.

Bando, S., & Nozawa, A. (2015). Detection of driver inattention from fluctuations in vehicle operating data. *Artificial Life and Robotics*, *20*(1), 28–33.

Banerjee, A., Datta, S., Pal, M., Konar, A., Tibarewala, D., & Janarthanan, R. (2013). Classifying electrooculogram to detect directional eye movements. *Procedia Technology*, *10*, 67–75.

Bang, J. W., Choi, J.-S., & Park, K. R. (2013). Noise reduction in brainwaves by using both EEG signals and frontal viewing camera images. *Sensors*, *13*(5), 6272–6294.

Banks, V. A., Plant, K. L., & Stanton, N. A. (2018). Driver error or designer error: Using the Perceptual Cycle Model to explore the circumstances surrounding the fatal Tesla crash on 7th May 2016. *Safety Science*, *108*, 278–285.

Banks, V. A., & Stanton, N. A. (2016). Keep the driver in control: Automating automobiles of the future. *Applied Ergonomics*, *53*, 389–395.

Barreto, A., Zhai, J., & Adjouadi, M. (2007). *Non-intrusive physiological monitoring for automated stress detection in human-computer interaction*. 29–38.

Bastuji, H., Perrin, F., & Garcia-Larrea, L. (2003). Event-related potentials during forced awakening: A tool for the study of acute sleep inertia. *Journal of Sleep Research*, *12*(3), 189–206.

Bekiaris, E. (1999). System for effective assessment of the driver state and vehicle control in emergency situations. *City. Project Report. Telematics TR1047 SAVE Telematics Consortium*. internal-pdf://219.125.180.76/systems for affective assesment of the driver.pdf

Berka, C., Levendowski, D. J., Cvetinovic, M. M., Petrovic, M. M., Davis, G., Lumicao, M. N., Zivkovic, V. T., Popovic, M. V., & Olmstead, R. (2004). Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset. *International Journal of Human-Computer Interaction*, *17*(2), 151–170.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, *78*(5), B231–B244.

Berntson, G. G., Quigley, K. S., Jang, J. F., & Boysen, S. T. (1990). An approach to artifact identification: Application to heart period data. *Psychophysiology*, *27*(5), 586–598.

Berntson, G. G., & Stowell, J. R. (1998). ECG artifacts and heart period variability: Don’t miss a beat! *Psychophysiology*, *35*(1), 127–132.

Berntson, G. G., Thomas Bigger Jr, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., Nagaraja, H. N., Porges, S. W., Saul, J. P., & Stone, P. H. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, *34*(6), 623–648.

Blatter, K., & Cajochen, C. (2007). Circadian rhythms in cognitive performance: Methodological constraints, protocols, theoretical underpinnings. *Physiology & Behavior*, *90*(2), 196–208.

Boersma, P., & Van Heuven, V. (2001). Speak and unSpeak with PRAAT. *Glot International*, *5*(9/10), 341–347.

Boksem, M. A., & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews*, *59*(1), 125–139.

Bonnet, J., Vai, F., Pioger, G., Garrigue, S., Clementy, J., & Cazeu, S. (1998). Circadian variations in minute ventilation can be reproduced by a pacemaker sensor. *Pacing and Clinical Electrophysiology*, *21*(4), 701–705.

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

Boyanov, B., & Hadjitodorov, S. (1997). Acoustic analysis of pathological voices. A voice analysis system for the screening of laryngeal diseases. *IEEE Engineering in Medicine and Biology Magazine*, *16*(4), 74–82.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology*, *49*(1), 1017–1034.

Brookhuis, K. A., de Vries, G., & De Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis & Prevention*, *23*(4), 309–316.

Brookhuis, K. A., & de Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis & Prevention*, *42*(3), 898–903.

Brouwer, W. H., & Withaar, F. K. (1997). Fitness to drive after traumatic brain injury. *Neuropsychological Rehabilitation*, *7*(3), 177–193.

Brown, E. N., Choe, Y., Luithardt, H., & Czeisler, C. A. (2000). A statistical model of the human core-temperature circadian rhythm. *American Journal of Physiology-Endocrinology and Metabolism*, *279*(3), E669–E683.

Brown, I. D. (1994). Driver fatigue. *Human Factors*, *36*(2), 298–314.

Bruck, D., & Pisani, D. L. (1999). The effects of sleep inertia on decision-making performance. *Journal of Sleep Research*, *8*(2), 95–103.

Bunce, S. C., Izzetoglu, K., Ayaz, H., Shewokis, P., Izzetoglu, M., Pourrezaei, K., & Onaral, B. (2011). *Implementation of fNIRS for monitoring levels of expertise and mental workload*. 13–22.

Bundele, M. M., & Banerjee, R. (2010). Design of Early Fatigue Detection Elements of a Wearable Computing System for the Prevention of Road Accidents. *Intelligent Systems and Applications (ISA), 2010 2nd International Workshop On*, 1–4. internal-pdf://119.132.3.39/design of early fatigue detection elements of.pdf

Bundele, M. M., & Banerjee, R. (2009). *Detection of fatigue of vehicular driver using skin conductance and oximetry pulse: A neural network approach*. 739–744.

Burbank, D. P., & Webster, J. G. (1978). Reducing skin potential motion artefact by skin abrasion. *Medical and Biological Engineering and Computing*, *16*(1), 31–38.

Burnham, J. F. (2006). Scopus database: A review. *Biomedical Digital Libraries*, *3*. https://doi.org/10.1186/1742-5581-3-1

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (2007). *Handbook of psychophysiology*. Cambridge University Press.

Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*, *89*(3–4), 319–325.

Caird, J. K. (2015). On the Effects of Listening and Talking to Humans and Devices on Driving. *Human Factors*, *57*(8), 1325–1327. https://doi.org/10.1177/0018720815575942

Cajochen, C, Wyatt, J. K., Czeisler, C. A., & Dijk, D. J. (2002). Separation of circadian and wake duration-dependent modulation of EEG activation during wakefulness. *Neuroscience*, *114*(4), 1047–1060. https://doi.org/10.1016/S0306-4522(02)00209-9

Cajochen, Christian, Khalsa, S. B. S., Wyatt, J. K., Czeisler, C. A., & Dijk, D.-J. (1999). EEG and ocular correlates of circadian melatonin phase and human performance decrements during sleep loss. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *277*(3), R640–R649.

Caldwell, J. A. (2005). Fatigue in aviation. *Travel Medicine and Infectious Disease*, *3*(2), 85–96.

Casner, S. M., Hutchins, E. L., & Norman, D. (2016). The challenges of partially automated driving. *Communications of the ACM*, *59*(5), 70–77.

Cavallari, J. M., Fang, S. C., Mittleman, M. A., & Christiani, D. C. (2010). Circadian variation of heart rate variability among welders. *Occupational and Environmental Medicine*, oem. 2010.055210.

Chakraborty, M., & Aoyon, A. N. H. (2014). Implementation of computer vision to detect driver fatigue or drowsiness to reduce the chances of vehicle accident. *Electrical Engineering and Information & Communication Technology (ICEEICT), 2014 International Conference On*, 1–5. internal-pdf://148.114.239.12/Implementation of Computer Vision to detect dr.pdf

Chan, M., Nyazika, S., & Singhal, A. (2016). Effects of a front-seat passenger on driver attention: An electrophysiological approach. *Transportation Research Part F: Traffic Psychology and Behaviour*, *43*, 67–79.

Chang, K.-M. (2010). Arrhythmia ECG noise reduction by ensemble empirical mode decomposition. *Sensors*, *10*(6), 6063–6080.

Chaplin, W. F., John, O. P., & Goldberg, L. R. (1988). Conceptions of states and traits: Dimensional attributes with ideals as prototypes. *Journal of Personality and Social Psychology*, *54*(4), 541.

Chellappa, S. L., Morris, C. J., & Scheer, F. A. J. L. (2019). Effects of circadian misalignment on cognition in chronic shift workers. *Scientific Reports*, *9*(1), 699.

Cheng, S.-Y., & Hsu, H.-T. (2011). Mental fatigue measurement using EEG. In *Risk management trends*. IntechOpen.

Chialastri, A. (2012). Automation in aviation. In *Automation*. InTech.

Chipman, M., & Jin, Y. L. (2009). Drowsy drivers: The effect of light and circadian rhythm on crash occurrence. *Safety Science*, *47*(10), 1364–1370. https://doi.org/10.1016/j.ssci.2009.03.005

Chong, J. W., Dao, D. K., Salehizadeh, S. M. A., McManus, D. D., Darling, C. E., Chon, K. H., & Mendelson, Y. (2014). Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection–reduction approach. Part I: motion and noise artifact detection. *Annals of Biomedical Engineering*, *42*(11), 2238–2250.

Chuang, C.-H., Huang, C.-S., Ko, L.-W., & Lin, C.-T. (2015). An EEG-based perceptual function integration network for application to drowsy driving. *Knowledge-Based Systems*, *80*, 143–152.

Cluydts, R., De Valck, E., Verstraeten, E., & Theys, P. (2002). Daytime sleepiness and its evaluation. *Sleep Medicine Reviews*, *6*(2), 83–96.

Cohen, B. H., Davidson, R. J., Senulis, J. A., Saron, C. D., & Weisman, D. R. (1992). Muscle tension patterns during auditory attention. *Biological Psychology*, *33*(2–3), 133–156.

Cuesta, M., Boudreau, P., Cermakian, N., & Boivin, D. B. (2017). Skin Temperature Rhythms in Humans Respond to Changes in the Timing of Sleep and Light. *Journal of Biological Rhythms*, *32*(3), 257–273.

Cummings, L., Dane, A., Rhodes, J., Lynch, P., & Hughes, A. M. (2000). Diurnal variation in the quantitative EEG in healthy adult volunteers. *British Journal of Clinical Pharmacology*, *50*(1), 21–26.

Cutmore, T. R. H., & James, D. A. (1999). Identifying and reducing noise in psychophysiological recordings. *International Journal of Psychophysiology*, *32*(2), 129–150.

da Silveira, T. L., Kozakevicius, A. J., & Rodrigues, C. R. (2016). Automated drowsiness detection through wavelet packet analysis of a single EEG channel. *Expert Systems with Applications*, *55*, 559–565.

Damousis, I. G., & Tzovaras, D. (2008). Fuzzy fusion of eyelid activity indicators for hypovigilance-related accident prediction. *IEEE Transactions on Intelligent Transportation Systems*, *9*(3), 491–500.

De Luca, C. J., Gilmore, L. D., Kuznetsov, M., & Roy, S. H. (2010). Filtering the surface EMG signal: Movement artifact and baseline noise contamination. *Journal of Biomechanics*, *43*(8), 1573–1579.

De Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 196–217.

Del Corral, P., Schurman, R. C., Kinza, S. S., Fitzgerald, M. J., Kordick, C. A., Rusch, J. L., & Nadolski, J. B. (2016). Salivary but not plasma cortisone tracks the plasma cortisol response to exercise: Effect of time of day. *Journal of Endocrinological Investigation*, *39*(3), 315–322.

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

Delorme, D., & Song, B. (2001). *Human driver model for SmartAHS*.

Dhupati, L. S., Kar, S., Rajaguru, A., & Routray, A. (2010). A novel drowsiness detection scheme based on speech analysis with validation using simultaneous EEG recordings. *Automation Science and Engineering (CASE), 2010 IEEE Conference On*, 917–921. internal-pdf://73.30.157.239/a novel drowsiness detection scheme based on s.pdf

Di Stasi, L. L., Antolí, A., Gea, M., & Cañas, J. J. (2011). A neuroergonomic approach to evaluating mental workload in hypermedia interactions. *International Journal of Industrial Ergonomics*, *41*(3), 298–304.

Di Stasi, L. L., McCamy, M. B., Pannasch, S., Renner, R., Catena, A., Cañas, J. J., Velichkovsky, B. M., & Martinez-Conde, S. (2015). Effects of driving time on microsaccadic dynamics. *Experimental Brain Research*, *233*(2), 599–605.

Dijk, D.-J. (1999). Circadian variation of EEG power spectra in NREM and REM sleep in humans: Dissociation from body temperature. *Journal of Sleep Research*, *8*(3), 189–195. https://doi.org/10.1046/j.1365-2869.1999.00159.x

Dijk, Duffy, & Czeisler. (1992). Circadian and sleep/wake dependent aspects of subjective alertness and cognitive performance. *Journal of Sleep Research*, *1*(2), 112–117.

Dijk, Shanahan, Duffy, Ronda, & Czeisler. (1997). Variation of electroencephalographic activity during non-rapid eye movement and rapid eye movement sleep with phase of circadian melatonin rhythm in humans. *The Journal of Physiology*, *505*(Pt 3), 851–858.

Dogan, E., Rahal, M.-C., Deborne, R., Delhomme, P., Kemeny, A., & Perrin, J. (2017). Transition of control in a partially automated vehicle: Effects of anticipation and non-driving-related task involvement. *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 205–215.

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

Douma, L. G., & Gumz, M. L. (2017). Circadian Clock-Mediated Regulation of Blood Pressure. *Free Radical Biology and Medicine*.

Drews, F. A., Yazdani, H., Godfrey, C. N., Cooper, J. M., & Strayer, D. L. (2009). Text messaging during simulated driving. *Human Factors*, *51*(5), 762–770.

Drury, C. G., Green, B. D., Chen, J., & Henry, E. L. (2006). Sleep, sleepiness, fatigue, and vigilance in a day and night inspection task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*, 66–70.

Duchowski, A. T., & Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice*. Springer.

Ebrahim, P., Stolzmann, W., & Yang, B. (2013). Eye movement detection for assessing driver drowsiness by electrooculography. *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference On*, 4142–4148. internal-pdf://255.149.69.45/Eye Movement Detection for Assessing Driver Dr.pdf

Ekhart, D., Wicht, H., Kersken, T., Ackermann, H., Kaczmarczyk, M., Pretzsch, G., Alexander, H., & Korf, H. W. (2018). Dynamics of core body temperature cycles in long-term measurements under real life conditions in women. *Chronobiology International, 35*(1), 8–23.

Endsley, M. R. (1996). Automation and situation awareness. *Automation and Human Performance: Theory and Applications*, *20*, 163–181.

Endsley, M. R. (2016). From Here to Autonomy: Lessons Learned From Human–Automation Research. *Human Factors*, *59*(1), 5–27. https://doi.org/10.1177/0018720816681350

Endsley, M. R. (2019). The limits of highly autonomous vehicles: An uncertain future. *Ergonomics*, *62*(4), 496–499. https://doi.org/10.1080/00140139.2019.1563330

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of cognitive load on driving performance: The cognitive control hypothesis. *Human Factors*, *59*(5), 734–764.

Eriksson, A., & Stanton, N. A. (2017). Takeover Time in Highly Automated Vehicles: Noncritical Transitions to and From Manual Control. *Human Factors*, *59*(4), 689–705.

Fairclough, S. H., & Houston, K. (2004). A metabolic measure of mental effort. *Biological Psychology*, *66*(2), 177–190.

Fan, X., Sun, Y., Yin, B., & Guo, X. (2010). Gabor-based dynamic representation for human fatigue monitoring in facial image sequences. *Pattern Recognition Letters*, *31*(3), 234–243.

Fathima, N., & Umarani, K. (2016). Reduction of noise in EEG signal using Faraday's cage and wavelets transform: A comparative study. *Int. J. Eng. Sci. & Comp*, *6*(7), 8566–8569.

Ferrara, M., & De Gennaro, L. (2000). The sleep inertia phenomenon during the sleep-wake transition: Theoretical and operational issues. *Aviation, Space, and Environmental Medicine*, *71*(8), 843–848.

Filtness, A J, Armstrong, K. A., Watson, A., & Smith, S. S. (2017). Sleep-related vehicle crashes on low speed roads. *Accident Analysis & Prevention*, *99*, 279–286. https://doi.org/10.1016/j.aap.2016.12.002

Filtness, Ashleigh J, Anund, A., Fors, C., Ahlström, C., Akerstedt, T., & Kecklund, G. (2014). Sleep-related eye symptoms and their potential for identifying driver sleepiness. *Journal of Sleep Research*, *23*(5), 568–575.

Frank, G., Halberg, F., Harner, R., Matthews, J., Johnson, E., Gravem, H., & Andrus, V. (1966). Circadian periodicity, adrenal corticosteroids, and the EEG of normal man. *Journal of Psychiatric Research*, *4*(2), 73–86.

Ftouni, S., Sletten, T. L., Howard, M., Anderson, C., Lenné, M. G., Lockley, S. W., & Rajaratnam, S. M. (2013). Objective and subjective measures of sleepiness, and their associations with on-road driving events in shift workers. *Journal of Sleep Research*, *22*(1), 58–69.

Fu, R., Wang, H., & Zhao, W. (2016). Dynamic driver fatigue detection using hidden Markov model in real driving condition. *Expert Systems with Applications*, *63*, 397–411.

G. Švec, J., S. Popolo, P., & R. Titze, I. (2003). Measurement of vocal doses in speech: Experimental procedure and signal processing. *Logopedics Phoniatrics Vocology*, *28*(4), 181–192.

Gander, P. H., Mulrine, H. M., van den Berg, M. J., Smith, A. A. T., Signal, T. L., Wu, L. J., & Belenky, G. (2015). Effects of sleep/wake history and circadian phase on proposed pilot fatigue safety performance indicators. *Journal of Sleep Research*, *24*(1), 110–119.

Garbarino, S., Nobili, L., Philip, P., Plazzi, G., Campus, C., Morrone, E., De Carli, F., & Gp, S. (2016). Circadian Sleep Propensity and Alcohol Interaction at the Wheel. *Journal of Clinical Sleep Medicine*, *12*(7), 1011–1017. https://doi.org/10.5664/jcsm.5938

Gauthier, A., Davenne, D., Martin, A., Cometti, G., & Hoecke, J. Van. (1996). Diurnal rhythm of the muscular performance of elbow flexors during isometric contractions. *Chronobiology International*, *13*(2), 135–146.

Gibbon, D., Moore, R., & Winski, R. (1997). *Handbook of standards and resources for spoken language systems*. Walter de Gruyter.

Gillberg, M., Kecklund, G., & Akerstedt, T. (1996). Sleepiness and performance of professional drivers in a truck simulator—Comparisons between day and night driving. *Journal of Sleep Research*, *5*(1), 12–15.

Giorgetti, R., Marcotulli, D., Tagliabracci, A., & Schifano, F. (2015). Effects of ketamine on psychomotor, sensory and cognitive functions relevant for driving ability.

Forensic Science International, *252*, 127–142. https://doi.org/10.1016/j.forsciint.2015.04.024

Gonçalves, M., Amici, R., Lucas, R., \AAkerstedt, T., Cirignotta, F., Horne, J., Léger, D., McNicholas, W. T., Partinen, M., Téran-Santos, J., Peigneux, P., & Grote, L. (2015). Sleepiness at the wheel across Europe: A survey of 19 countries. *Journal of Sleep Research*, *24*(3), 242–253. https://doi.org/10.1111/jsr.12267

Grace, R., Byrne, V. E., Bierman, D. M., Legrand, J.-M., Gricourt, D., Davis, B., Staszewski, J. J., & Carnahan, B. (1998). *A drowsy driver detection system for heavy vehicles*. *2*, I36-1.

Graw, P., Kräuchi, K., Knoblauch, V., Wirz-Justice, A., & Cajochen, C. (2004). Circadian and wake-dependent modulation of fastest and slowest reaction times during the psychomotor vigilance task. *Physiology and Behavior*, *80*(5), 695–701. https://doi.org/10.1016/j.physbeh.2003.12.004

Gubin, D. G., Weinert, D., Rybina, S. V, Danilova, L. A., Solovieva, S. V, Durov, A. M., Prokopiev, N. Y., & Ushakov, P. A. (2017). Activity, sleep and ambient light have a different impact on circadian blood pressure, heart rate and body temperature rhythms. *Chronobiology International*, *34*(5), 632–649.

Gundel, A., & Witthöft, H. (1983). Circadian rhythm in the EEG of man. *International Journal of Neuroscience*, *19*(1–4), 287–292.

Ha, U., & Yoo, H.-J. (2016). A multimodal drowsiness monitoring ear-module system with closed-loop real-time alarm. *Biomedical Circuits and Systems Conference (BioCAS), 2016 IEEE*, 536–539. internal-pdf://66.64.59.207/A Multimodal Drowsiness Monitoring Ear-Module.pdf

Haigney, D. E., Taylor, R. G., & Westerman, S. J. (2000). Concurrent mobile (cellular) phone use and driving performance: Task demand characteristics and compensatory processes. *Transportation Research Part F: Traffic Psychology and Behaviour*, *3*(3), 113–121.

Häkkänen, H., & Summala, H. (2001). Fatal traffic accidents among trailer truck drivers and accident causes as viewed by other truck drivers. *Accident Analysis & Prevention*, *33*(2), 187–196.

Hancock, P. A. (2019). Some pitfalls in the promises of automated and autonomous vehicles. *Ergonomics*, *62*(4), 479–495. https://doi.org/10.1080/00140139.2018.1498136

Hancock, P. A. (2015). Automobility: The coming use of fully-automated on-road vehicles. *Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2015 IEEE International Inter-Disciplinary Conference On*, 137–139. internal-pdf://0855911177/Automobility. The comin use of fully automated.pdf

Haq, Z. A., & Hasan, Z. (2016). Eye-blink rate detection for fatigue determination. *Information Processing (IICIP), 2016 1st India International Conference On*, 1–5. internal-pdf://0761265581/Eye-blink rate detection for fatigue determina.pdf

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139–183). Elsevier.

Haynes, J. M. (2007). The ear as an alternative site for a pulse oximeter finger clip sensor. *Respiratory Care*, *52*(6), 727–729.

He, J., Becic, E., Lee, Y.-C., & McCarley, J. S. (2011). Mind wandering behind the wheel: Performance and oculomotor correlates. *Human Factors*, *53*(1), 13–21.

He, J., Choi, W., Yang, Y., Lu, J., Wu, X., & Peng, K. (2017). Detection of driver drowsiness using wearable devices: A feasibility study of the proximity sensor. *Applied Ergonomics*, *65*, 473–480.

Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, *6*(2), 156–166.

Heaney, J. L. J., Phillips, A. C., & Carroll, D. (2012). Aging, health behaviors, and the diurnal rhythm and awakening response of salivary cortisol. *Experimental Aging Research*, *38*(3), 295–314.

Heath, H., & Cowley, S. (2004a). Developing a grounded theory approach: A comparison of Glaser and Strauss. *International Journal of Nursing Studies*, *41*(2), 141–150.

Heath, H., & Cowley, S. (2004b). Developing a grounded theory approach: A comparison of Glaser and Strauss. *International Journal of Nursing Studies*, *41*(2), 141–150.

Heikoop, D. D., de Winter, J. C. F., van Arem, B., & Stanton, N. A. (2016). Psychological constructs in driving automation: A consensus model and critical comment on construct proliferation. *Theoretical Issues in Ergonomics Science*, *17*(3), 284–303.

Hermida, R. C., Ayala, D. E., & Iglesias, M. (2004). Circadian blood pressure variability as a function of parity in normotensive pregnant women. *The Journal of Clinical Hypertension*, *6*(3), 126–133.

Higgins, J. S., Michael, J., Austin, R., \AAkerstedt, T., Van Dongen, H. P., Watson, N., Czeisler, C., Pack, A. I., & Rosekind, M. R. (2017). Asleep at the Wheel-The Road to Addressing Drowsy Driving. *Sleep*, *40*(2). https://doi.org/10.1093/sleep/zsx001

Higuchi, S., Liu, Y., Yuasa, T., Maeda, A., & Motohashi, Y. (2000). Diurnal variation in the P300 component of human cognitive event-related potential. *Chronobiology International*, *17*(5), 669–678.

Hoel, J., Jaffard, M., Boujon, C., & Van Elslande, P. (2011). Different forms of attentional disturbances involved in driving accidents. *IET Intelligent Transport Systems*, *5*(2), 120–126.

Hofer-Tinguely, G., Achermann, P., Landolt, H.-P., Regel, S. J., Rétey, J. V., Dürr, R., Borbély, A. A., & Gottselig, J. M. (2005). Sleep inertia: Performance changes after sleep, rest and active waking. *Cognitive Brain Research*, *22*(3), 323–331. https://doi.org/10.1016/j.cogbrainres.2004.09.013

Hogervorst, M. A., Brouwer, A.-M., & Van Erp, J. B. (2014). Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload. *Frontiers in Neuroscience*, *8*, 322.

Horne, James A, & Baulk, S. D. (2004). Awareness of sleepiness when driving. *Psychophysiology*, *41*(1), 161–165.

Horne, Jim A, & Reyner, L. A. (1995). Sleep related vehicle accidents. *Bmj*, *310*(6979), 565–567.

Horrey, W. J., & Wickens, C. D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors*, *48*(1), 196–205.

Hosking, S. G., Young, K. L., & Regan, M. A. (2009). The effects of text messaging on young drivers. *Human Factors*, *51*(4), 582–592.

Hot, P., Naveteur, J., Leconte, P., & Sequeira, H. (1999). Diurnal variations of tonic electrodermal activity. *International Journal of Psychophysiology*, *33*(3), 223–230.

Howard, M. E., Jackson, M. L., Berlowitz, D., O'Donoghue, F., Swann, P., Westlake, J., Wilkinson, V., & Pierce, R. J. (2014). Specific sleepiness symptoms are indicators of performance impairment during sleep deprivation. *Accident Analysis & Prevention*, *62*, 1–8.

Hunter, E. J., Spielman, J., Starr, A., & Popolo, P. (1997). Acoustic Voice Recording,"I am seeking recommendations for voice recording. *Analysis*, *40*(4), 894–899.

Hwang, S. H., Seo, S., Yoon, H. N., Baek, H. J., Cho, J., Choi, J. W., Lee, Y. J., Jeong, D.-U., & Park, K. S. (2015). Sleep period time estimation based on electrodermal activity. *IEEE Journal of Biomedical and Health Informatics*, *21*(1), 115–122.

Hwang, T., Kim, M., Hong, S., & Park, K. S. (2016). *Driver drowsiness detection using the in-ear EEG*. 4646–4649.

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2006). Subjective sleepiness, simulated driving performance and blink duration: Examining individual differences. *Journal of Sleep Research*, *15*(1), 47–53.

Islam, M. K., Rastegarnia, A., & Yang, Z. (2016). Methods for artifact detection and removal from scalp EEG: A review. *Neurophysiologie Clinique/Clinical Neurophysiology*, *46*(4–5), 287–305.

Itoh, M. (2009). *Individual differences in effects of secondary cognitive activity during driving on temperature at the nose tip*. 7–11.

Jackson, M. L., Kennedy, G. A., Clarke, C., Gullo, M., Swann, P., Downey, L. A., Hayley, A. C., Pierce, R. J., & Howard, M. E. (2016). The utility of automated measures of ocular metrics for detecting driver drowsiness during extended wakefulness. *Accident Analysis & Prevention*, *87*, 127–133.

Jackson, M. L., Raj, S., Croft, R. J., Hayley, A. C., Downey, L. A., Kennedy, G. A., & Howard, M. E. (2016). Slow eyelid closure as a measure of driver drowsiness and its relationship to performance. *Traffic Injury Prevention*, *17*(3), 251–257.

Jainta, S., & Baccino, T. (2010). Analyzing the pupil response due to increased cognitive demand: An independent component analysis study. *International Journal of Psychophysiology*, *77*(1), 1–7.

Jap, B. T., Lal, S., Fischer, P., & Bekiaris, E. (2009). Using EEG spectral components to assess algorithms for detecting fatigue. *Expert Systems with Applications*, *36*(2), 2352–2359.

Ji, Q., & Yang, X. (2002). Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, *8*(5), 357–377.

Jiao, Y., & Lu, B.-L. (2016). Detecting slow eye movement for recognizing driver's sleep onset period with EEG features. *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of The*, 4658–4661. internal-pdf://88.134.3.254/detecting slowe eye movement for recognizing d.pdf

Jödicke, J., Olbrich, S., Sander, C., Minkwitz, J., Chittka, T., Himmerich, H., & Hegerl, U. (2013). Separation of low-voltage EEG-activity during mental activation from that during transition to drowsiness. *Brain Topography*, *26*(4), 538–546.

Johannsdottir, K. R., & Herdman, C. M. (2010). The Role of Working Memory in Supporting Drivers' Situation Awareness for Surrounding Traffic. *Human Factors*, *52*(6), 663–673. https://doi.org/10.1177/0018720810385427

Johns, M. W. (2000). A sleep physiologist's view of the drowsy driver. *Transportation Research Part F: Traffic Psychology and Behaviour*, *3*(4), 241–249. https://doi.org/10.1016/S1369-8478(01)00008-0

Kaber, D. B. (2018). A conceptual framework of autonomous and automated agents. *Theoretical Issues in Ergonomics Science*, *19*(4), 406–430. https://doi.org/10.1080/1463922X.2017.1363314

Kaduk, S. I., Roberts, A. P., & Stanton, N. A. (2020). The manual shift in phase: The impact of circadian phase on semi-autonomous driving. What can we learn from current understanding in manual driving? *Theoretical Issues in Ergonomics Science*, 1–21.

Kaduk S.I., Roberts A.P., Stanton N.A. (2021) Psychophysiological Predictors of Motion Sickness in the Driving Simulator. In: Cassenti D., Scataglini S., Rajulu S., Wright J. (eds) Advances in Simulation and Digital Human Modeling. AHFE 2020. Advances in Intelligent Systems and Computing, vol 1206. Springer, Cham

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology*, *117*(7), 1574–1581.

Kamzanova, A. T., Kustubayeva, A. M., & Matthews, G. (2014). Use of EEG workload indices for diagnostic monitoring of vigilance decrement. *Human Factors*, *56*(6), 1136–1149.

Kartsch, V., Benatti, S., Rossi, D., & Benini, L. (2017). *A wearable EEG-based drowsiness detection system with blink duration and alpha waves analysis*. 251–254.

Kawanaka, H., Miyaji, M., Bhuiyan, M., & Oguri, K. (2013). Identification of cognitive distraction using physiological features for adaptive driving safety supporting system. *International Journal of Vehicular Technology*, *2013*.

Keith, K., Trentacoste, M. F., Depue, L., Granda, T. M., Huckaby, E., Ibarguen, B., Kantowitz, B. H., Lum, W., & Wilson, T. (2005). *Roadway human factors and behavioral safety in Europe*. United States. Federal Highway Administration.

Khan, M. J., & Hong, K.-S. (2015). Passive BCI based on drowsiness detection: An fNIRS study. *Biomedical Optics Express*, *6*(10), 4063–4078.

Khan, M. J., Liu, X., Bhutta, M. R., & Hong, K.-S. (2016). *Drowsiness detection using fNIRS in different time windows for a passive BCI*. 227–231.

Kim, J., Hong, J., Kim, N., Cha, E., & Lee, T.-S. (2007). *Two algorithms for detecting respiratory rate from ECG signal*. 4069–4071.

Kirst, M., Glauner, B., & Ottenbacher, J. (2011). Using DWT for ECG motion artifact reduction with noise-correlating signals. *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 4804–4807.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data*.

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New England Journal of Medicine*, *370*(1), 54–59.

Kolff, M., Hofman, W. F., Kerkhof, G. A., van den Broek, P., Coenen, A., Ruigt, G. S. F., van Bemmel, A. L., de Boer, T., Hofman, W., & van Luijtelaar, G. (2003). Electroencephalographic features of sleep inertia. *Sleep/Wake Research in the Netherlands*, 99–105.

Kosmadopoulos, A., Sargent, C., Zhou, X., Darwent, D., Matthews, R. W., Dawson, D., & Roach, G. D. (2017). The efficacy of objective and subjective predictors of driving performance during sleep restriction and circadian misalignment. *Accident Analysis & Prevention*, *99*, 445–451. https://doi.org/10.1016/j.aap.2015.10.014

Krajewski, J., Batliner, A., & Golz, M. (2009). Acoustic sleepiness detection: Framework and validation of a speech-adapted pattern recognition approach. *Behavior Research Methods*, *41*(3), 795–804.

Krajewski, J., & Kröger, B. (2007). *Using prosodic and spectral characteristics for sleepiness detection*. Eighth Annual Conference of the International Speech Communication Association.

Krajewski, J., Schnieder, S., Sommer, D., Batliner, A., & Schuller, B. (2012). Applying multiple classifiers and non-linear dynamics features for detecting sleepiness from speech. *Neurocomputing*, *84*, 65–75.

Krajewski, J., Trutschel, U., Golz, M., Sommer, D., & Edwards, D. (2009). *Estimating fatigue from predetermined speech samples transmitted by operator communication systems*.

Kristjansson, S. D., Stern, J. A., Brown, T. B., & Rohrbaugh, J. W. (2009). Detecting phasic lapses in alertness using pupillometric measures. *Applied Ergonomics*, *40*(6), 978–986.

Kumari, B. M. K., & Kumar, P. R. (2017). A survey on drowsy driver detection system. *Big Data Analytics and Computational Intelligence (ICBDAC), 2017 International Conference On*, 272–279. internal-pdf://251.211.206.0/zza survey on drowsy driver detection system.pdf

Kwai, P. C. A., Imtiaz, S. A., Bowyer, S., & Rodriguez-Villegas, E. (2016). An algorithm for automatic detection of drowsiness for use in wearable EEG systems. *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of The*, 3523–3526. internal-pdf://206.158.50.148/An algorithm for automatic detection of drowsi.pdf

Kyriakidis, M., de Winter, J. C. F., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., Martens, M. H., Bengler, K., Andersson, J., Merat, N., Reed, N., Flament, M., Hagenzieker, M., & Happee, R. (2019). A human factors perspective on automated driving. *Theoretical Issues in Ergonomics Science*, 1–27. https://doi.org/10.1080/1463922X.2017.1293187

Lal, S. K., & Craig, A. (2001). Electroencephalography activity associated with driver fatigue: Implications for a fatigue countermeasure device. *Journal of Psychophysiology*, *15*(3), 183.

Lal, S. K., & Craig, A. (2002). Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiology*, *39*(3), 313–321.

Lal, S. K., & Craig, A. (2005). Reproducibility of the spectral components of the electroencephalogram during driver fatigue. *International Journal of Psychophysiology*, *55*(2), 137–143.

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

Lal, S. K. L., & Craig, A. (2002). Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiology*, *39*(3), 313–321.

Lambert, F. (2016). Understanding the fatal Tesla accident on autopilot and the NHTSA probe. *Electrek, July*.

Lamble, D., Kauranen, T., Laakso, M., & Summala, H. (1999). Cognitive load and detection thresholds in car following situations: Safety implications for using mobile (cellular) telephones while driving. *Accident Analysis & Prevention*, *31*(6), 617–623.

Lanatà, A., Scilingo, E. P., Nardini, E., Loriga, G., Paradiso, R., & De-Rossi, D. (2009). Comparative evaluation of susceptibility to motion artifact in different wearable systems for monitoring respiratory rate. *IEEE Transactions on Information Technology in Biomedicine*, *14*(2), 378–386.

Larkin, K. T., Semenchuk, E. M., Frazer, N. L., Suchday, S., & Taylor, R. L. (1998). Cardiovascular and behavioral response to social confrontation: Measuring real-life stress in the laboratory. *Annals of Behavioral Medicine*, *20*(4), 294–301.

Lazar, A. S., Lazar, Z. I., & Dijk, D.-J. (2015). Circadian regulation of slow waves in human sleep: Topographical aspects. *NeuroImage*, *116*, 123–134. https://doi.org/10.1016/j.neuroimage.2015.05.012

Leger, D., Philip, P., Jarriault, P., Metlaine, A., & Choudat, D. (2009). Effects of a combination of napping and bright light pulses on shift workers' sleepiness at the wheel: A pilot study. *Journal of Sleep Research*, *18*(4), 472–479.

Lemmer, B. (1995). Clinical chronopharmacology: The importance of time in drug treatment. *Ciba Foundation Symposium*, *183*, 235–253.

Leng, L. B., Giin, L. B., & Chung, W.-Y. (2015). *Wearable driver drowsiness detection system based on biomedical and motion sensors*. 1–4.

Lenné, M. G., Triggs, T. J., & Redman, J. R. (1998). Interactive Effects of Sleep Deprivation, Time of Day, and Driving Experience on a Driving Task. *Sleep*, *21*(1), 38–44. https://doi.org/10.1093/sleep/21.1.38

Li, G., Lee, B.-L., & Chung, W.-Y. (2015). Smartwatch-based wearable EEG system for driver drowsiness detection. *IEEE Sensors Journal*, *15*(12), 7169–7180.

Li, P., Jiang, W., & Su, F. (2016). *Single-channel EEG-based mental fatigue detection based on deep belief network*. 367–370.

Li, T., Zhao, Y., Sun, Y., Gao, Y., Su, Y., Hetian, Y., & Chen, M. (2015). Near-infrared spectroscopy assessment of divided visual attention task-invoked cerebral hemodynamics during prolonged true driving. *SPIE BiOS*, 93052I–93052I–7. internal-pdf://0560981046/fNIRS.pdf

Liang, Y., & Lee, J. D. (2010). Combining cognitive and visual distraction: Less than the sum of its parts. *Accident Analysis & Prevention*, *42*(3), 881–890.

Lin, C.-T., Ko, L.-W., Chung, I.-F., Huang, T.-Y., Chen, Y.-C., Jung, T.-P., & Liang, S.-F. (2006). Adaptive EEG-based alertness estimation system by using ICA-based fuzzy neural networks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, *53*(11), 2469–2476.

Lin, C.-T., Pal, N. R., Chuang, C.-Y., Jung, T.-P., Ko, L.-W., & Liang, S.-F. (2008). *An EEG-based subject-and session-independent drowsiness detection*. 3448–3454.

Linhart, P., Ratcliffe, V. F., Reby, D., & Špinka, M. (2015). Expression of emotional arousal in two different piglet call types. *PloS One*, *10*(8), e0135414.

Liu, C. C., Hosking, S. G., & Lenné, M. G. (2009). Predicting driver drowsiness using vehicle measures: Recent insights and future challenges. *Journal of Safety Research*, *40*(4), 239–245.

Liu, G., Wu, D., Mei, Z., Zhu, Q., & Wang, L. (2013). Automatic detection of respiratory rate from electrocardiogram, respiration induced plethysmography and 3D acceleration signals. *Journal of Central South University*, *20*(9), 2423–2431.

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.

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.

Liu, Y., & Ozguner, U. (2007). *Human driver model and driver decision making for intersection driving*. 642–647.

Lopez, A., Ferrero, F. J., Valledor, M., Campo, J. C., & Postolache, O. (2016). A study on electrode placement in EOG systems for medical applications. *Medical Measurements and Applications (MeMeA), 2016 IEEE International Symposium On*, 1–5.

Lowden, A., Anund, A., Kecklund, G., Peters, B., & Akerstedt, T. (2009). Wakefulness in young and elderly subjects driving at night in a car simulator. *Accident Analysis & Prevention*, *41*(5), 1001–1007. https://doi.org/10.1016/j.aap.2009.05.014

Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human–human and human–automation trust: An integrative review. *Theoretical Issues in Ergonomics Science*, *8*(4), 277–301. https://doi.org/10.1080/14639220500337708

Maglione, A., Borghini, G., Arico, P., Borgia, F., Graziani, I., Colosimo, A., Kong, W., Vecchiato, G., & Babiloni, F. (2014). Evaluation of the workload and drowsiness during car driving by using high resolution EEG activity and neurophysiologic indices. *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 6238–6241. internal-pdf://224.134.50.165/Evaluation of the workload and drowsiness duri.pdf

Marinescu, A. C., Sharples, S., Ritchie, A. C., Sanchez Lopez, T., McDowell, M., & Morvan, H. P. (2018). Physiological parameter response to variation of mental workload. *Human Factors*, *60*(1), 31–56.

Márquez, L., Cantillo, V., & Arellana, J. (2015). Mobile phone use while driving: A hybrid modeling approach. *Accident Analysis & Prevention*, *78*, 73–80. https://doi.org/10.1016/j.aap.2015.02.016

Maryn, Y., De Bodt, M., & Roy, N. (2010). The Acoustic Voice Quality Index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders*, *43*(3), 161–174.

Mathewson, K. E., Harrison, T. J. L., & Kizuk, S. A. D. (2017). High and dry? Comparing active dry EEG electrodes to active and passive wet electrodes. *Psychophysiology*, *54*(1), 74–82.

Ma'touq, J., Al-Nabulsi, J., Al-Kazwini, A., Baniyassien, A., Al-Haj Issa, G., & Mohammad, H. (2014). Eye blinking-based method for detecting driver drowsiness. *Journal of Medical Engineering & Technology*, *38*(8), 416–419.

Matthews, G. (2002). Towards a transactional ergonomics for driver stress and fatigue. *Theoretical Issues in Ergonomics Science*, *3*(2), 195–211. https://doi.org/10.1080/14639220210124120

Matthews, G., & Desmond, P. A. (2002). Task-induced fatigue states and simulated driving performance. *The Quarterly Journal of Experimental Psychology: Section A*, *55*(2), 659–686.

Matthews, R. W., Ferguson, S. A., Zhou, X., Kosmadopoulos, A., Kennaway, D. J., & Roach, G. D. (2012). Simulated driving under the influence of extended wake, time of day and sleep restriction. *Accident Analysis & Prevention*, *45*, 55–61. https://doi.org/10.1016/j.aap.2011.09.027

Matthews, R. W., Ferguson, S. A., Zhou, X., Sargent, C., Darwent, D., Kennaway, D. J., & Roach, G. D. (2012). Time-of-Day Mediates the Influences of Extended Wake and Sleep Restriction on Simulated Driving. *Chronobiology International*, *29*(5), 572–579. https://doi.org/10.3109/07420528.2012.675845

Matthews, R. W., Ferguson, S. A., Zhou, X., Sargent, C., Darwent, D., Kennaway, D. J., & Roach, G. D. (2012). Time-of-day mediates the influences of extended wake and sleep restriction on simulated driving. *Chronobiology International*, *29*(5), 572–579.

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(3), 218–224. https://doi.org/10.1016/j.trf.2008.11.005

May, J. G., Kennedy, R. S., Williams, M. C., Dunlap, W. P., & Brannan, J. R. (1990). Eye movement indices of mental workload. *Acta Psychologica*, *75*(1), 75–89.

May, J. G., Kennedy, R. S., Williams, M. C., Dunlap, W. P., & Brannan, J. R. (1990). Eye movement indices of mental workload. *Acta Psychologica*, *75*(1), 75–89.

McKenna, P. (1998). Fitnessto drive: A neuropsychological perspective. *Journal of Mental Health*, *7*(1), 9–18.

Measures, S. for P. R. A. H. C. on E., Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034.

Mehta, R. K., & Parasuraman, R. (2013). Neuroergonomics: A review of applications to physical and cognitive work. *Frontiers in Human Neuroscience*, *7*, 889.

Melnicuk, V., Birrell, S., Crundall, E., & Jennings, P. (2016). Towards hybrid driver state monitoring: Review, future perspectives and the role of consumer electronics. *Intelligent Vehicles Symposium (IV), 2016 IEEE*, 1392–1397. internal-pdf://233.210.5.168/3- Towards Hybrid Driver State Monitoring.pdf

Merat, N., Jamson, A. H., Lai, F. C., Daly, M., & Carsten, O. M. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 274–282.

Merriam-Webster. (2018). *Dictionary*.

Michael, L., Passmann, S., & Becker, R. (2012). Electrodermal lability as an indicator for subjective sleepiness during total sleep deprivation. *Journal of Sleep Research*, *21*(4), 470–478.

Milosevic, S. (1997). Drivers' fatigue studies. *Ergonomics*, *40*(3), 381–389.

Minhad, K. N., Ali, S. H. M., & Reaz, M. B. I. (2017). Happy-anger emotions classifications from electrocardiogram signal for automobile driving safety and awareness. *Journal of Transport & Health*, *7*, 75–89. https://doi.org/10.1016/j.jth.2017.11.001

Ministers to agree new global road safety agenda to 2030. (n.d.). Retrieved 27 March 2020, from https://www.who.int/news-room/detail/07-02-2020-ministers-to-agree-new-global-road-safety-agenda-to-2030

Mitler, M. M., Carskadon, M. A., Czeisier, C. A., Dement, W. C., Dinges, D. F., & Graeber, R. C. (1988). Catastrophes, Sleep, and Public Policy: Consensus Report. *Sleep*, *11*(1), 100–109. https://doi.org/10.1093/sleep/11.1.100

Mittal, A., Kumar, K., Dhamija, S., & Kaur, M. (2016). Head movement-based driver drowsiness detection: A review of state-of-art techniques. *Engineering and Technology (ICETECH), 2016 IEEE International Conference On*, 903–908. internal-pdf://254.24.94.145/zHead movement-based driver drowsiness detecti.pdf

Miyaji, M., Kawanaka, H., & Oguri, K. (2009). Driver's cognitive distraction detection using physiological features by the adaboost. *Intelligent Transportation Systems, 2009. ITSC'09. 12th International IEEE Conference On*, 1–6. internal-pdf://0283854263/Drivers cognitive distraction detection using.pdf

Miyake, S., Yamada, S., Shoji, T., Takae, Y., Kuge, N., & Yamamura, T. (2009). Physiological responses to workload change. A test/retest examination. *Applied Ergonomics*, *40*(6), 987–996.

Mongia, P. K., & Sharma, R. (2014). Estimation and statistical analysis of human voice parameters to investigate the influence of psychological stress and to determine the vocal tract transfer function of an individual. *Journal of Computer Networks and Communications*, *2014*.

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

Mullaney, D., Kripke, D., & Messin, S. (1980). Wrist-actigraphic estimation of sleep time. *Sleep*, *3*(1), 83–92.

Murata, A, Naitoh, K., & Karwowski, W. (2017.). A method for predicting the risk of virtual crashes in a simulated driving task using behavioural and subjective drowsiness measures. *Ergonomics*, *60*(5), 714–730. https://doi.org/10.1080/00140139.2016.1223885

Murata, Atsuo, Fujii, Y., & Naitoh, K. (2015). Multinomial logistic regression model for predicting driver's drowsiness using behavioral measures. *Procedia Manufacturing*, *3*, 2426–2433.

Murata, Atsuo, & Hiramatsu, Y. (2008). *Evaluation of drowsiness by HRV measures-basic study for drowsy driver detection*. *2008*(1), 99–102.

Murata, Atsuo, Naitoh, K., & Karwowski, W. (2017). A method for predicting the risk of virtual crashes in a simulated driving task using behavioural and subjective drowsiness measures. *Ergonomics*, *60*(5), 714–730.

Muzet, A., Nicolas, A., Tassi, P., Dewasmes, G., & Bonneau, A. (1995). Implementation of napping in industry and the problem of sleep inertia. *Journal of Sleep Research*, *4*, 67–69.

Nater, U. M., Rohleder, N., Schlotz, W., Ehlert, U., & Kirschbaum, C. (2007). Determinants of the diurnal course of salivary alpha-amylase. *Psychoneuroendocrinology*, *32*(4), 392–401.

Neumann, P. G. (2016). Risks of automation: A cautionary total-system perspective of our cyberfuture. *Communications of the ACM*, *59*(10), 26–30.

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods*, *192*(1), 152–162.

Noureddin, B., Lawrence, P. D., & Birch, G. E. (2012). Online removal of eye movement and blink EEG artifacts using a high-speed eye tracker. *IEEE Transactions on Biomedical Engineering*, *59*(8), 2103–2110.

O'Connor, S. S., Shain, L. M., Whitehill, J. M., & Ebel, B. E. (2017). Measuring a conceptual model of the relationship between compulsive cell phone use, in-vehicle cell phone use, and motor vehicle crash. *Accident Analysis & Prevention*, *99*, 372–378. https://doi.org/10.1016/j.aap.2016.12.016

Ogorevc, J., Podlesek, A., Geršak, G., & Drnovšek, J. (2011). The effect of mental stress on psychophysiological parameters. *Medical Measurements and Applications Proceedings (MeMeA), 2011 IEEE International Workshop On*, 294–299. internal-pdf://148.165.174.193/The effect of mental stress on psychophysiolog.pdf

Oken, B S, Salinsky, M. C., & Elsas, S. M. (2006). Vigilance, alertness, or sustained attention: Physiological basis and measurement. *Clinical Neurophysiology*, *117*(9), 1885–1901.

O'Regan, S., Faul, S., & Marnane, W. (2013). Automatic detection of EEG artefacts arising from head movements using EEG and gyroscope signals. *Medical Engineering & Physics*, *35*(7), 867–874.

Otmani, S., Rogé, J., & Muzet, A. (2005). Sleepiness in professional drivers: Effect of age and time of day. *Accident Analysis & Prevention*, *37*(5), 930–937. https://doi.org/10.1016/j.aap.2005.04.011

Pack, A. I., Pack, A. M., Rodgman, E., Cucchiara, A., Dinges, D. F., & Schwab, C. W. (1995). Characteristics of crashes attributed to the driver having fallen asleep. *Accident Analysis & Prevention*, *27*(6), 769–775. https://doi.org/10.1016/0001-4575(95)00034-8

Pajcin, M., Banks, S., White, J. M., Dorrian, J., Paech, G. M., Grant, C., Johnson, K., Tooley, K., Fidock, J., & Kamimori, G. H. (2017). Decreased salivary alpha-amylase levels are associated with performance deficits during sleep loss. *Psychoneuroendocrinology*, *78*, 131–141.

Palinko, O., Kun, A. L., Shyrokov, A., & Heeman, P. (2010). Estimating cognitive load using remote eye tracking in a driving simulator. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 141–144. internal-pdf://0881045796/Estimating cognitive load using remote eye tra.pdf

Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P. D., Chouvarda, I., Bekiaris, E., & Maglaveras, N. (2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clinical Neurophysiology*, *118*(9), 1906–1922.

Parasuraman, R. (2011). Neuroergonomics: Brain, cognition, and performance at work. *Current Directions in Psychological Science*, *20*(3), 181–186.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 140–160.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 140–160.

Parker, D., Reason, J. T., Manstead, A. S. R., & Stradling, S. G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics*, *38*(5), 1036–1048.

Parnell, K. J., Stanton, N. A., & Plant, K. L. (2016). Exploring the mechanisms of distraction from in-vehicle technology: The development of the PARRC model. *Safety Science*, *87*, 25–37.

Patrick, K. C., Imtiaz, S. A., Bowyer, S., & Rodriguez-Villegas, E. (2016). *An algorithm for automatic detection of drowsiness for use in wearable EEG systems*. 3523–3526.

Perala, C. H., & Sterling, B. S. (2007). *Galvanic skin response as a measure of soldier stress*. Army Research Lab Aberdeen Proving Ground Md Human Research and Engineering ….

Perrier, J., Jongen, S., Vuurman, E., Bocca, M., Ramaekers, J., & Vermeeren, A. (2016). Driving performance and EEG fluctuations during on-the-road driving following sleep deprivation. *Biological Psychology*, *121*, 1–11.

Petridou, E., & Moustaki, M. (2000). Human factors in the causation of road traffic crashes. *European Journal of Epidemiology*, *16*(9), 819–826.

Phinyomark, A., Phukpattaranont, P., & Limsakul, C. (2012). The usefulness of wavelet transform to reduce noise in the SEMG signal. *EMG Methods for Evaluating Muscle and Nerve Function*, 107–132.

Phipps-Nelson, J. O., Redman, J. R., & Rajaratnam, S. M. W. (2011). Temporal profile of prolonged, night-time driving performance: Breaks from driving temporarily reduce time-on-task fatigue but not sleepiness. *Journal of Sleep Research*, *20*(3), 404–415.

Pickering, T. G., Schwartz, J. E., & Stone, A. (1996). Behavioral influences on diurnal blood pressure rhythms. *Annals of the New York Academy of Sciences*, *783*(1), 132–140.

Pinegger, A., Wriessnegger, S. C., Faller, J., & Müller-Putz, G. R. (2016). Evaluation of different EEG acquisition systems concerning their suitability for building a brain–computer interface: Case studies. *Frontiers in Neuroscience*, *10*, 441.

Plichta, B. (2002). Best practices in the acquisition, processing, and analysis of acoustic speech signals. *University of Pennsylvania Working Papers in Linguistics*, *8*(3), 16.

Plöchl, M., Ossandón, J. P., & König, P. (2012). Combining EEG and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in Human Neuroscience*, *6*, 278.

Pope, C. N., Bell, T. R., & Stavrinos, D. (2017). Mechanisms behind distracted driving behavior: The role of age and executive function in the engagement of distracted driving. *Accident Analysis & Prevention*, *98*, 123–129. https://doi.org/10.1016/j.aap.2016.09.030

Popieul, J. C., Simon, P., & Loslever, P. (2003). *Using driver's head movements evolution as a drowsiness indicator*. 616–621.

Posada-Quintero, H. F., Bolkhovsky, J. B., Reljin, N., & Chon, K. H. (2017). Sleep deprivation in young and healthy subjects is more sensitively identified by higher frequencies of electrodermal activity than by skin conductance level evaluated in the time domain. *Frontiers in Physiology*, *8*, 409.

Prattichizzo, F. A., & Galetta, F. (1995). Circadian changes of ambulatory blood pressure in an unselected population. *Archives of Gerontology and Geriatrics*, *20*(2), 185–191.

Puspasari, M. A., Iridiastadi, H., & Sutalaksana, I. Z. (2015). Oculomotor response in measuring driving fatigue. *Industrial Engineering and Engineering Management (IEEM), 2015 IEEE International Conference On*, 1091–1095. internal-pdf://0296502213/Oculomotor response in measuring driving fatig.pdf

Qian, D., Wang, B., Qing, X., Zhang, T., Zhang, Y., Wang, X., & Nakamura, M. (2016). Bayesian nonnegative CP decomposition-based feature extraction algorithm for drowsiness detection. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(8), 1297–1308.

Rafferty, L. A., Stanton, N. A., & Walker, G. H. (2010). The famous five factors in teamwork: A case study of fratricide. *Ergonomics*, *53*(10), 1187–1204.

Rahman, A., Sirshar, M., & Khan, A. (2015). Real time drowsiness detection using eye blink monitoring. *Software Engineering Conference (NSEC), 2015 National*, 1–7. internal-pdf://158.132.110.75/Real Time Drowsiness Detection using Eye Blink.pdf

Rahman, M. Z. U., Shaik, R. A., & Reddy, D. R. K. (2011). Efficient sign based normalized adaptive filtering techniques for cancelation of artifacts in ECG signals: Application to wireless biotelemetry. *Signal Processing*, *91*(2), 225–239.

Ram, M. R., Madhav, K. V., Krishna, E. H., Komalla, N. R., & Reddy, K. A. (2012). A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter. *IEEE Transactions on Instrumentation and Measurement*, *61*(5), 1445–1457.

Ranzijn, R., & Lack, L. (1997). The pupillary light reflex cannot be used to measure sleepiness. *Psychophysiology*, *34*(1), 17–22.

Reddy, M. S., Chander, K. P., & Rao, K. S. (2011). Complex wavelet transform driven de-noising method for an EOG signals. *India Conference (INDICON), 2011 Annual IEEE*, 1–5.

Redelmeier, D. A., & Tibshirani, R. J. (1997). Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine*, *336*(7), 453–458.

Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis & Prevention*, *43*(5), 1771–1781.

Resalat, S. N., Saba, V., & Afdideh, F. (2012). *A novel system for driver's sleepiness detection using SSVEP*. 343–347.

Richter, P., Wagner, T., Heger, R., & Weise, G. (1998). Psychophysiological analysis of mental load during driving on rural roads-a quasi-experimental field study. *Ergonomics*, *41*(5), 593–609.

Roberts, A. P., Stanton, N. A., Plant, K. L., Fay, D. T., & Pope, K. A. (2020). You say it is physical, I say it is functional; let us call the whole thing off! Simulation: An application divided by lack of common language. *Theoretical Issues in Ergonomics Science*, *21*(5), 507–536.

Rodríguez-Ibáñez, N., García-González, M. A., Fernández-Chimeno, M., & Ramos-Castro, J. (2011). Drowsiness detection by thoracic effort signal analysis in real driving environments. *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 6055–6058. internal-pdf://137.247.8.140/Drowsiness detection by thoracic effort signal.pdf

Roscoe, A. H. (1992). Assessing pilot workload. Why measure heart rate, HRV and respiration? *Biological Psychology*, *34*(2), 259–287.

Ross, V., Jongen, E. M. M., Wang, W., Brijs, T., Brijs, K., Ruiter, R. A. C., & Wets, G. (2014). Investigating the influence of working memory capacity when driving behavior is combined with cognitive load: An LCT study of young novice drivers. *Accident Analysis & Prevention*, *62*, 377–387. https://doi.org/10.1016/j.aap.2013.06.032

Rumschlag, G., Palumbo, T., Martin, A., Head, D., George, R., & Commissaris, R. L. (2015). The effects of texting on driving performance in a driving simulator: The influence of driver age. *Accident Analysis & Prevention*, *74*, 145–149. https://doi.org/10.1016/j.aap.2014.10.009

Sagaspe, P., Taillard, J., Amiéva, H., Beck, A., Rascol, O., Dartigues, J.-F., Capelli, A., & Philip, P. (2012). Influence of age, circadian and homeostatic processes on inhibitory motor control: A Go/Nogo task study. *PloS One*, *7*(6), e39410.

Sagberg, F. (1999). Road accidents caused by drivers falling asleep. *Accident Analysis & Prevention*, *31*(6), 639–649.

Sahayadhas, A., Sundaraj, K., & Murugappan, M. (2013). Drowsiness detection during different times of day using multiple features. *Australasian Physical & Engineering Sciences in Medicine*, *36*(2), 243–250.

Sahayadhas, A., Sundaraj, K., Murugappan, M., & Palaniappan, R. (2015a). A physiological measures-based method for detecting inattention in drivers using machine learning approach. *Biocybernetics and Biomedical Engineering*, *35*(3), 198–205.

Sahayadhas, A., Sundaraj, K., Murugappan, M., & Palaniappan, R. (2015b). Physiological signal based detection of driver hypovigilance using higher order spectra. *Expert Systems with Applications*, *42*(22), 8669–8677.

Salimetrics, L., & SalivaBio, L. (2011). Saliva collection and handling advice. *State College, PA: Salimetrics LLC, SalivaBio LLC*, 1–14.

Salmon, P., Stanton, N., Walker, G., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics*, *37*(2), 225–238.

Samn, S. W., & Perelli, L. P. (1982). *Estimating aircrew fatigue: A technique with application to airlift operations*. School of Aerospace Medicine Brooks Afb tx.

Sandberg, D., Anund, A., Fors, C., Kecklund, G., Karlsson, J. G., Wahde, M., & \AAkerstedt, T. (2011). The characteristics of sleepiness during real driving at night—A study of driving performance, physiology and subjective experience. *Sleep*, *34*(10), 1317–1325. https://doi.org/10.5665/SLEEP.1270

Sanpeng, D., Xiaoli, X., Xuecui, Y., & Dehua, M. (2010). *Research on the driver fatigue monitoring method based on the Dempster-Shafer theory*. 4176–4179.

Saritha, C., Sukanya, V., & Murthy, Y. N. (2008). ECG signal analysis using wavelet transforms. *Bulg. J. Phys*, *35*(1), 68–77.

Schleicher, R., Galley, N., Briest, S., & Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: Looking tired? *Ergonomics*, *51*(7), 982–1010.

Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., & Othersen, I. (2015). The interaction between highly automated driving and the development of drowsiness. *Procedia Manufacturing*, *3*, 6652–6659.

Schreinicke, G., Hinz, A., Kratzsch, J., Hüber, B., & Voigt, G. (1990). Stress-related changes of saliva cortisol in VDU operators. *International Archives of Occupational and Environmental Health*, *62*(4), 319–321.

Seiler, S. J. (2015). Hand on the wheel, mind on the mobile: An analysis of social factors contributing to texting while driving. *Cyberpsychology, Behavior, and Social Networking*, *18*(2), 72–78. https://doi.org/10.1089/cyber.2014.0535

Sharma, M. K., & Bundele, M. M. (2015). *Design & analysis of k-means algorithm for cognitive fatigue detection in vehicular driver using oximetry pulse signal*. 1–6.

Shin, D., Sakai, H., & Uchiyama, Y. (2011). Slow eye movement detection can prevent sleep-related accidents effectively in a simulated driving task. *Journal of Sleep Research*, *20*(3), 416–424.

Shultz, M. (2007). Comparing test searches in PubMed and Google Scholar. *Journal of the Medical Library Association*, *95*(4), 442–445. https://doi.org/10.3163/1536-5050.95.4.442

Siddiqui, U., & Shaikh, A. N. (2013). An overview of electrooculography. *International Journal of Advanced Research in Computer and Communication Engineering*, *2*(11), 4328–4330.

Simon, M., Schmidt, E. A., Kincses, W. E., Fritzsche, M., Bruns, A., Aufmuth, C., Bogdan, M., Rosenstiel, W., & Schrauf, M. (2011). EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Clinical Neurophysiology*, *122*(6), 1168–1178.

Smith, P., Shah, M., & da Vitoria Lobo, N. (2003). Determining driver visual attention with one camera. *IEEE Transactions on Intelligent Transportation Systems*, *4*(4), 205–218.

Sneddon, A., Mearns, K., & Flin, R. (2013). Stress, fatigue, situation awareness and safety in offshore drilling crews. *Safety Science, 56*, 80–88. https://doi.org/10.1016/j.ssci.2012.05.027

Solaz, J., Laparra-Hernández, J., Bande, D., Rodríguez, N., Veleff, S., Gerpe, J., & Medina, E. (2016). Drowsiness detection based on the analysis of breathing rate obtained from real-time image recognition. *Transportation Research Procedia*, *14*, 3867–3876.

Sriraam, N., Shri, T. K. P., & Maheshwari, U. (2016). Recognition of wake-sleep stage 1 multichannel eeg patterns using spectral entropy features for drowsiness detection. *Australasian Physical & Engineering Sciences in Medicine*, *39*(3), 797–806.

Stanton, N. A. (2015). Responses to autonomous vehicles. *Ingenia*, *62*(69).

Stanton, N. A., & Marsden, P. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science*, *24*(1), 35–49.

Stanton, N. A., & Salmon, P. M. (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. *Safety Science*, *47*(2), 227–237.

Stanton, N. A., Young, M., & McCaulder, B. (1997). Drive-by-wire: The case of driver workload and reclaiming control with adaptive cruise control. *Safety Science*, *27*(2–3), 149–159.

Stanton, N. A., & Young, M. S. (1998). Vehicle automation and driving performance. *Ergonomics*, *41*(7), 1014–1028.

Stanton, N. A., & Young, M. S. (2000). A proposed psychological model of driving automation. *Theoretical Issues in Ergonomics Science*, *1*(4), 315–331.

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: A possible measure of fatigue. *Human Factors*, *36*(2), 285–297.

Stolz, G., Aschoff, J. C., Born, J., & Aschoff, J. (1988). VEP, physiological and psychological circadian variations in humans. *Journal of Neurology*, *235*(5), 308–313.

Strahler, J., Berndt, C., Kirschbaum, C., & Rohleder, N. (2010). Aging diurnal rhythms and chronic stress: Distinct alteration of diurnal rhythmicity of salivary α-amylase and cortisol. *Biological Psychology*, *84*(2), 248–256. https://doi.org/10.1016/j.biopsycho.2010.01.019

Strayer, D. L., & Drew, F. A. (2004). Profiles in driver distraction: Effects of cell phone conversations on younger and older drivers. *Human Factors*, *46*(4), 640–649.

Subasi, A. (2005). Application of classical and model-based spectral methods to describe the state of alertness in EEG. *Journal of Medical Systems*, *29*(5), 473–486.

Suda, M., Fukuda, M., Sato, T., Iwata, S., Song, M., Kameyama, M., & Mikuni, M. (2009). Subjective feeling of psychological fatigue is related to decreased reactivity in ventrolateral prefrontal cortex. *Brain Research*, *1252*, 152–160.

Sweeney, K. T., Ward, T. E., & McLoone, S. F. (2012). Artifact removal in physiological signals—Practices and possibilities. *IEEE Transactions on Information Technology in Biomedicine*, *16*(3), 488–500.

Tan, X., Uchiyama, M., Shibui, K., Tagaya, H., Suzuki, H., Kamei, Y., Kim, K., Aritaka, S., Ozaki, A., & Takahashi, K. (2003). Circadian rhythms in humans' delta sleep electroencephalogram. *Neuroscience Letters*, *344*(3), 205–208.

Tassi, P., Bonnefond, A., Engasser, O., Hoeft, A., Eschenlauer, R., & Muzet, A. (2006). EEG spectral power and cognitive performance during sleep inertia: The effect of normal sleep duration and partial sleep deprivation. *Physiology & Behavior*, *87*(1), 177–184.

Tassi, P., & Muzet, A. (2000). Sleep inertia. *Sleep Medicine Reviews*, *4*(4), 341–353. https://doi.org/10.1053/smrv.2000.0098

Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., & Picard, R. (2015). Automatic identification of artifacts in electrodermal activity data. *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 1934–1937.

Tran, Y., Thuraisingham, R., Wijesuriya, N., Craig, A., & Nguyen, H. (2014). Using S-transform in EEG analysis for measuring an alert versus mental fatigue state. *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 5880–5883. internal-pdf://117.187.213.161/Using S-transform in EEG analysis for measurin.pdf

Tsukahara, K., Moriwaki, S., Hotta, M., Fujimura, T., & Kitahara, T. (2004). A study of diurnal variation in wrinkles on the human face. *Archives of Dermatological Research*, *296*(4), 169–174.

Tuunainen, A., Kripke, D. F., Cress, A. C., & Youngstedt, S. D. (2001). Retinal circadian rhythms in humans. *Chronobiology International*, *18*(6), 957–971.

Urigüen, J. A., & Garcia-Zapirain, B. (2015). EEG artifact removal—State-of-the-art and guidelines. *Journal of Neural Engineering*, *12*(3), 031001.

Van Boxtel, A. (2001). Optimal signal bandwidth for the recording of surface EMG activity of facial, jaw, oral, and neck muscles. *Psychophysiology*, *38*(1), 22–34.

Van Boxtel, A. (2010). *Facial EMG as a tool for inferring affective states*. 104–108.

van den Berg, M. J., Wu, L. J., & Gander, P. H. (2016). Subjective Measurements of In-Flight Sleep, Circadian Variation, and Their Relationship with Fatigue. *Aerospace Medicine and Human Performance*, *87*(10), 869–875.

Van Dongen, H. P. A., & Dinges, D. F. (2000). Circadian rhythms in fatigue, alertness, and performance. *Principles and Practice of Sleep Medicine*, *20*, 391–399.

van Lieshout, P. (2003). *PRAAT*. https://web.stanford.edu/dept/linguistics/corpora/material/PRAAT_workshop_manual_v421.pdf, retrieved online 15.12.2020, 11:15.

Van Winsum, W. (2000). Measuring drowsiness and impairment in car driving. *Proceedings Vision in Vehicles*, *8*. internal-pdf://133.217.17.251/Measuring drowsiness and impairment in car dri.pdf

Veltman, J. A., & Gaillard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology*, *42*(3), 323–342.

Vicente, J., Laguna, P., Bartra, A., & Bailón, R. (2016). Drowsiness detection using heart rate variability. *Medical & Biological Engineering & Computing*, *54*(6), 927–937.

Vitabile, S., De Paola, A., & Sorbello, F. (2010). *Bright pupil detection in an embedded, real-time drowsiness monitoring system*. 661–668.

Vural, E., Cetin, M., Ercil, A., Littlewort, G., Bartlett, M., & Movellan, J. (2007). *Drowsy driver detection through facial movement analysis*. 6–18.

Wang, J., Sun, S., Fang, S., Fu, T., & Stipancic, J. (2017). Predicting drowsy driving in real-time situations: Using an advanced driving simulator, accelerated failure time model, and virtual location-based services. *Accident Analysis & Prevention*, *99*, 321–329.

Wang, M., Jeong, N., Kim, K., Choi, S., Yang, S., You, S., Lee, J., & Suh, M. (2016). Drowsy behavior detection based on driving information. *International Journal of Automotive Technology*, *17*(1), 165–173.

Wang, X., & Xu, C. (2016). Driver drowsiness detection based on non-intrusive metrics considering individual specifics. *Accident Analysis & Prevention*, *95*, 350–357.

Wang, Y., Xin, M., Bai, H., & Zhao, Y. (2017). Can variations in visual behavior measures be good predictors of driver sleepiness? A real driving test study. *Traffic Injury Prevention*, *18*(2), 132–138.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, *50*(3), 433–441.

Weisgerber, D. M., Nikol, M., & Mistlberger, R. E. (2017). Driving home from the night shift: A bright light intervention study. *Sleep Medicine*, *30*, 171–179. https://doi.org/10.1016/j.sleep.2016.09.010

Wesensten, N. J., & Badia, P. (1992). Time of day and semantic category effects on late components of the visual ERP. *Biological Psychology*, *33*(2–3), 173–193.

Wesley, A., Shastri, D., & Pavlidis, I. (2010). A novel method to monitor driver's distractions. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 4273–4278).

Whitmore, J., & Fisher, S. (1996a). Speech during sustained operations. *Speech Communication*, *20*(1–2), 55–70.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology and human performance*. Psychology Press.

Wiegmann, D. A., Rich, A., & Zhang, H. (2001). Automated diagnostic aids: The effects of aid reliability on users' trust and reliance. *Theoretical Issues in Ergonomics Science*, *2*(4), 352–367. https://doi.org/10.1080/14639220110110306

Wierwille, W. W., Wreggit, S. S., Kirn, C. L., Ellsworth, L. A., & Fairbanks, R. J. (1994). *Research on vehicle-based driver status/performance monitoring; development, validation, and refinement of algorithms for detection of driver drowsiness. Final report*. internal-pdf://187.124.170.83/Research on vehicle based driver status perfor.pdf

Wijesuriya, N., Tran, Y., & Craig, A. (2007). The psychophysiological determinants of fatigue. *International Journal of Psychophysiology*, *63*(1), 77–86.

Wilhelm, B., Bittner, E., Hofmann, A., Koerner, A., Peters, T., Lüdtke, H., & Wilhelm, H. (2015). Short-term reproducibility and variability of the pupillographic sleepiness test. *American Journal of Human Biology*, *27*(6), 862–866.

Wilkinson, R T, & Stretton, M. (1971). Performance after awakening at different times of night. *Psychonomic Science*, *23*(4 LB-Wilkinson1971), 283–285. https://doi.org/10.3758/bf03336113

.

Williams, A. F. (1985). Nighttime driving and fatal crash involvement of teenagers. *Accident Analysis & Prevention*, *17*(1), 1–5. https://doi.org/10.1016/0001-4575(85)90002-8

Williamson, A. M., & Friswell, R. (2008). Fatigue and driving: Disentangling the relative effects of time of day and sleep deprivation. *Proceedings Of*. internal-pdf://162.20.70.20/Fatigue_and_driving_Disentangling_the_relative.pdf

Wilson, G. F. (2002). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *The International Journal of Aviation Psychology*, *12*(1), 3–18.

Wilson, G. F., & Russell, C. A. (2003). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors*, *45*(4), 635–644.

World Health Organization. (2015). *Global status report on road safety 2015*. World Health Organization. internal-pdf://242.59.168.68/WHO global status report on road safety.pdf

World Health Organization. (2018). *Global status report on road safety 2018: Summary*. World Health Organization.

Wörle, J., Metz, B., Othersen, I., & Baumann, M. (2020). Sleep in highly automated driving: Takeover performance after waking up. *Accident Analysis & Prevention*, *144*, 105617.

Yang, J. H., & Jeong, H. B. (2015). Validity analysis of vehicle and physiological data for detecting driver drowsiness, distraction, and workload. *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference On*, 1238–1243. internal-pdf://84.111.33.59/zValidity Analysis of Vehicle and Physiologica.pdf

Yeo, M. V., Li, X., Shen, K., & Wilder-Smith, E. P. (2009). Can SVM be used for automatic EEG detection of drowsiness during car driving? *Safety Science*, *47*(1), 115–124.

Yoon, T.-J., Zhuang, X., Cole, J., & Hasegawa-Johnson, M. (2006). *Voice quality dependent speech recognition*. International Symposium on Linguistic Patterns in Spontaneous Speech.

Young, M. S., & Stanton, N. A. (2002). Attention and automation: New perspectives on mental underload and performance. *Theoretical Issues in Ergonomics Science*, *3*(2), 178–194.

Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors*, *44*(3), 365–375.

Young, M. S., & Stanton, N. A. (2007). What's skill got to do with it? Vehicle automation and driver mental workload. *Ergonomics*, *50*(8), 1324–1339.

Young, & Stanton. (2005). Mental workload. *Handbook of Human Factors and Ergonomics Methods*, 31–39.

Zeier, H., Brauchli, P., & Joller-Jemelka, H. I. (1996). Effects of work demands on immunoglobulin A and cortisol in air traffic controllers. *Biological Psychology*, *42*(3), 413–423.

Zhang, H., Wang, Z., Li, Y., Zhang, Q., & Zhang, J. (2017). Research on time relevant variables based fatigue level prediction model. *Transportation Information and Safety (ICTIS), 2017 4th International Conference On*, 1–5.

Zhu, X., Zheng, W.-L., Lu, B.-L., Chen, X., Chen, S., & Wang, C. (2014). *EOG-based drowsiness detection using convolutional neural networks*. 128-131