PREDICTION OF DRIVERS' PERFORMANCE IN HIGHLY AUTOMATED VEHICLES

By

MOHAMED TAHER ALREFAIE

School of Business and Economics

LOUGHBOROUGH UNIVERSITY

A thesis submitted to Loughborough University in accordance with the requirements of the degree of
DOCTOR OF PHILOSOPHY.

March 2019

ABSTRACT

Purpose: The aim of this research was to assess the predictability of driver's response to critical hazards during the transition from automated to manual driving in highly automated vehicles using their physiological data.

Method: A driving simulator experiment was conducted to collect drivers' physiological data before, during and after the transition from automated to manual driving. A total of 33 participants between 20 and 30 years old were recruited. Participants went through a driving scenario under the influence of different non-driving related tasks. The repeated measures approach was used to assess the effect of repeatability on the driver's physiological data. Statistical and machine learning methods were used to assess the predictability of drivers' response quality based on their physiological data collected before responding to a critical hazard.

Findings: - The results showed that the observed physiological data that was gathered before the transition formed strong indicators of the drivers' ability to respond successfully to a potential hazard after the transition. In addition, physiological behaviour was influenced by driver's secondary tasks engagement and correlated with the driver's subjective measures to the difficulty of the task. The study proposes new quality measures to assess the driver's response to critical hazards in highly automated driving. Machine learning results showed that response time is predictable using regression methods. In addition, the classification methods were able to classify drivers into low, medium and high-risk groups based on their quality measures values.

Research Implications: Proposed models help increase the safety of automated driving systems by providing insights into the drivers' ability to respond to future critical hazards. More research is required to find the influence of age, drivers' experience of the automated vehicles and traffic density on the stability of the proposed models.

Originality: The main contribution to knowledge of this study is the feasibility of predicting drivers' ability to respond to critical hazards using the physiological behavioural data collected before the transition from automated to manual driving. With the findings, automation systems could change the transition time based on the driver's physiological state to allow for the safest transition possible. In addition, it provides an insight into driver's readiness and therefore, allows the automated system to adopt the correct driving strategy and plan to enhance drivers experience and make the transition phase safer for everyone.

To the souls of my father and my grandfather whom I lost whilst writing this thesis.

ACKNOWLEDGEMENT

Academic: I'd like to thank my academic supervisors, Tom Jackson and Steve Summerskills for their efforts to get me through this journey. I'd like to thank Andrew Morris for his critical questions that greatly influenced the work in this thesis.

Professional: I'd like to thank Carmen Torres-Sanchez and Paul Conway for the exceptional support given to me and the rest of the PhD students at the Centre of Doctoral Training of Embedded Intelligence. I'm also grateful for the help of Donna Palmers who put her efforts and dedications to ensure the success of the CDT-EI cohort.

Personal: Without the continuous emotional support of my mother, I wouldn't have been able to be where I'm. This journey wouldn't have been possible without the encouragement and dedication my dad has given me throughout the years. I'm very grateful for the years I spent with my grandfather who taught me the fundamentals of everything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

The thesis has used a number of abbreviations throughout the chapters. Although the abbreviations were explained at their first instance in the text, they are listed here for the the reader's ease of reference.

AI	Artificial Intelligence
FRESH	Feature extraction based on scalable hypothesis tests
HAD	Highly Automated Driving
HMI	Human Machine Interface
HR	Heart rate
ML	Machine Learning
MTF	Markov Transition Field
NDR	Non-driving-related
NHTSA	National Highway Traffic Safety Administration
PD	Pupil Diameter
PerAngle	Mean percentage change od vehicle lateral speed
PERCLOS	Percentage of eye closure
PerSpeed	Mean percentage change of vehicle's longitudinal speed
RF	Random Forest
RM	Repeated measures
RT	Response time
SAE	Society of Automotive Engineers
SVM	Support Vector Machine
TOR	Takeover request
TQT	Twenty Questions Tasks
TTC	Time to Collision
VIF	Variance Inflation Factor
ADAS	Advanced driver-assistance systems
ACC	Adaptive Cruise Control
EEG	Electroencephalogram
ECG	Electrocardiogram
PPG	photoplethysmogram
TOPS	Take-over performance score
SuRT	Surrogate reference task

1

CHAPTER

Introduction

A “man in freedom” as Aristotle defined is the ultimate peak of human existence. The person in freedom is a person that has a complete personal agency but wholly liberated from any concern for the necessities of life. With the promise of artificial intelligence (AI) replacing humans’ repetitive work, humanity comes a step closer to Aristotle’s vision (Wolcott, 2018).

In the past few decades, vehicle automation has gained substantial traction in both industry and automotive research (Lu and Winter, 2015; Merat et al., 2012; NHTSA, 2013). In a race to achieve full automation, several manufacturers, technology start-ups and automotive leaders introduced different automation systems to handle several driving tasks — for example, Tesla Motors’ lateral and longitudinal control of their vehicles (Ingle and Phute, 2016). Several other manufacturers introduced technologies such as motorway steering wheel assistance (Volvo, 2013) in addition to other companies committing to bringing the first fully automated vehicles to the mass market (Welch and Behrmann, 2018).

Vehicle automation’s main benefits go further than just freeing humans from driving. The full potential of vehicle automation unlocks the possibilities to a new world of mobility. For example, it enables the optimisation of the road network which will maximise the traffic flow and capacity (Papageorgiou et al., 2015). The outcome of the automation opens new possibilities for disrupting the future of mobility and transportation. This automation, in turn, will have a substantial economic and environmental impact by increasing number of shared vehicles (Fagnant and Kockelman, 2013), reduction in carbon dioxide footprint and reduction in energy consumption (Anderson et al., 2016). In addition, full automation will reduce fatalities by eliminating the human error that contributes to almost 93% of road accidents (Sabey and Taylor, 1980). However, there are many challenges imposed by automation. First, full driving automation will not take place immediately. Instead, it will take a gradual increment until fully autonomous vehicles are achieved (Anderson et al., 2016).

During this period, some argue that fatalities may be caused by semi-automated vehicles (Louw, 2017). Initially, human errors arise due to poor human-system interaction (Reason, 1990). Semi-automated vehicles, in turn, are characterised as a joint cognitive system between the human and the machine (Bibby et al., 1975) and the interaction between the two entities may expand problems rather than solve them (Bainbridge, 1983). Primarily, the transition from automated to manual driving is thought to be the most critical point of human-machine interaction (Anderson et al., 2016; Merat et al., 2012). This could be due to poorly designed interfaces (Reuschenbach et al., 2010), lack of understanding between the system and the machine (Koo et al., 2015) and human inability to handle critical hazards due to their lack of situational awareness (Merat et al., 2012). Though, as mentioned earlier, one of the primary motivations for full automation is the reduction of fatalities. The paradox, however, is that in order to achieve full automation, more accidents may happen until the full automation is achieved (Louw, 2017).

With such concern, human factors of automated driving have gained substantial traction to help identify the detrimental effects of semi-automated vehicles (Hs, 2014; Saffarian et al., 2012). To manage the aforementioned issues, several studies proposed solutions to manage the human-machine interaction. For example, interfaces were designed to communicate the automated system's state (Bazilinskyy et al., 2018), provide shared haptic control (Abbink et al., 2012) and multimodal warning alerts (Bazilinskyy et al., 2018). Even though an enhanced level of communication was observed, those studies focused primarily on one side of the interaction; the machine.

For a successful human-machine interaction, the machine (i.e., the automated system) will need to understand the human's mental state and its potential effect on their driving performance. This includes the driver's ability to identify and handle critical incidents during the transition from automation to manual driving. With the automated system's ability to understand human's state, systems could plan to ensure a safer transition based on driver's readiness.

1.1 Automation

The Human Condition, written by historian and philosopher Hannah Arendt, introduced a comprehensive framework for understanding human work in western history (Arendt and Canovan, 1998). The *Vita Activa,* as defined by her, consists of three levels; *Labour*, *Work* and *Action*.

"Labour generates metabolic necessities — the inputs, such as food, that sustain human life. Work creates the physical artefacts and infrastructure that define our world, and often outlast us — from homes and goods to works of art. Action encompasses interactive, communicative activities between human beings — the public sphere. In action, we explore and assert our distinctiveness as human beings and seek immortality.", (Wolcott, 2018)

Reflecting *Vita Activa* on driving, automation helps eliminate the driving labour, freeing humans to focus on long-lasting activities of both *Work* and *Action*. Full automation means, vehicles will perform sensing, reasoning and control. However, current automated systems require human's collaboration to achieve their primary aim of driving. This form of reliance requires transparency (Lyons, 2013) and open communication between the two major entities of the driving process; the human and the machine to form a *shared distributed situation awareness* (Stanton et al., 2006).

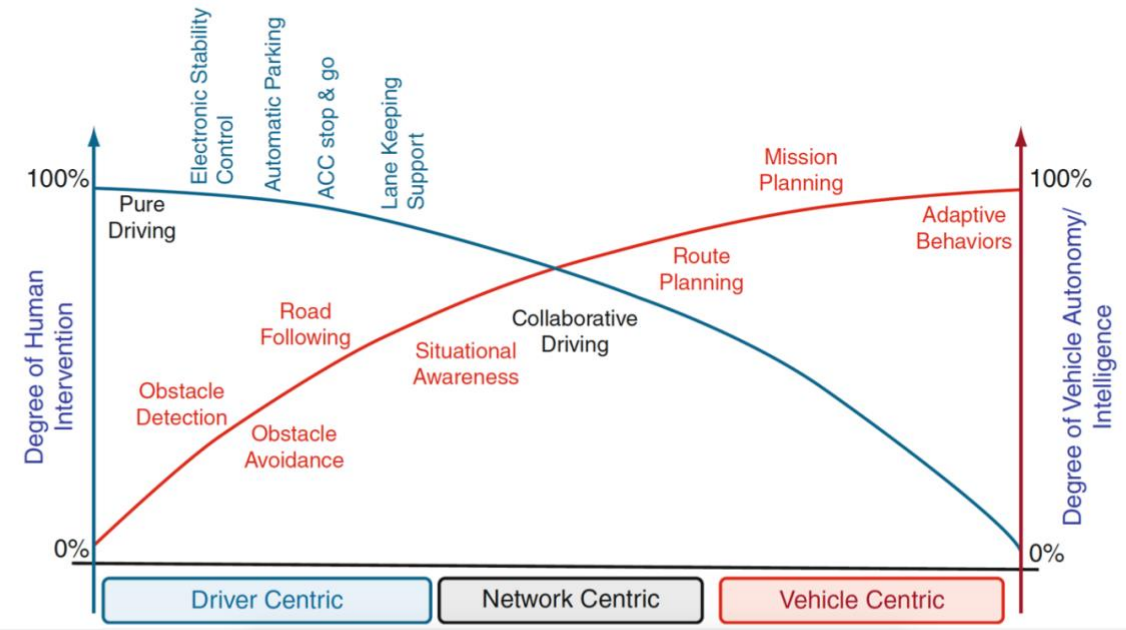


Figure 1: Impact of automation on driver's activities of the driving task. (Ibañez-Guzmán et al., 2012)

This collaboration is seen by drivers as a supervisory role where they can communicate their main goals or instructions which in turn are executed by the automated systems (Flemisch et al., 2012). This form of perception encourages drivers to be reliant on the automated system. However, the 'reliance' level of drivers is highly dependent on the intelligence degree of the automated system, the driver's experience (Larsson et al., 2014) and complacency (Parasuraman and Manzey, 2010). Figure 1 shows the driver's intervention decreasing as the automated system's abilities increase. The intersecting point between the two lines is identified as the collaborative driving that defines the point when drivers become more operator than an interactive driver of the vehicle. In the following section, automated driving and its impact on drivers are discussed, and the levels of automation are explained in detail.

1.2 Automated Driving

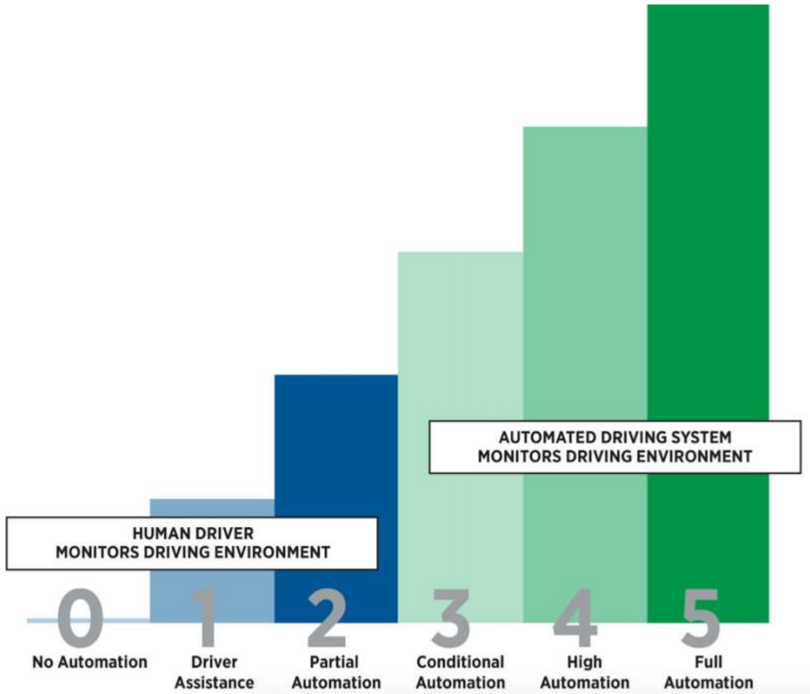
The concept of automated vehicles was first introduced in experiments in the early 1920s (Sentinel, 1926). New York Times article brought the light to the concept (New York Times, 1925) and since then, several manufacturers vision was directed towards the idea. Japan's Tsukuba Mechanical Engineering Laboratory developed the first concept that uses signal processing on an analogue computer to direct the vehicle to direct itself based on the white street markings (Weber, 2014). Old prototypes were heavily dependent on pre-existing infrastructure for guidance and navigation. Later in the 1980s and afterwards, vehicles were thought to operate independently of any infrastructure by relying on sensing hardware and software (Weber, 2014).

The United States Defence Advanced Research Projects Agency (DARPA) triggered the race among manufacturers and researchers to take self-driving vehicles from concept to realisation (Weber, 2014). This was done through a series of competitions called *Grand Challenge* 2004 and 2005 (Buehler et al., 2007) in addition to *Urban Challenge* in 2007 (Buehler et al., 2009). Several vehicles raced through suburban and urban areas using automated driving systems. Soon after, several autonomous vehicles were tested on roads such as Google Car (Poczter, SL & Jankovic, 2014). Since then, the race among top tech companies and start-ups started to create safe and reliable self-driving vehicles (Welch and Behrmann, 2018).

Researchers, regulators and policymakers put significant efforts in creating the necessary taxonomies to manage the degrees of vehicle automation. Two organisations, National Highway Traffic Safety Administration (NHTSA, 2013) and Society of Automotive Engineers (SAE International, 2018, 2016, 2014) developed two frameworks independently for classifying different levels of driving automation based on the tasks performed by humans and the automated systems. Both frameworks are similar; however, SAE's framework assumes the system is capable of monitoring the driving environment without necessarily being activated (SAE International, 2018). Both frameworks were highly criticised for establishing discretised levels of automation which in reality will not be the case since the systems will evolve naturally and gradually (Inagaki and Sheridan, 2018). Additionally, SAE’s model is the most widely cited model in the literature and has been revised twice since the SAE’s first report. Therefore, the SAE framework is defined and discussed in this study.

1.3 Different Levels of Automated Driving

As these systems evolve, different levels of automation have been characterised based on the system’s ability to intervene in longitudinal and lateral control of the vehicle (SAE International, 2018) as illustrated in Figure 2. Automation levels 1 and 2 have been achieved by several car manufacturers as a function of their advanced driving systems. In both levels, drivers are the primary agents of the driving task. However, the automated system is the central controller of the driving in Level 3 and above while drivers are not forced to monitor the driving environment (SAE International, 2018). Though, the driver’s duty is still to monitor the process while these systems have longitudinal and lateral control over the vehicle.



SAE level	Name	Narrative Definition	Execution of Steering and Acceleration/ Deceleration	*Monitoring* of Driving Environment	Fallback Performance of *Dynamic Driving Task*	System Capability *(Driving Modes)*
***Human driver* monitors the driving environment**						
0	No Automation	the full-time performance by the *human driver* of all aspects of the *dynamic driving task*, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
1	Driver Assistance	the *driving mode*-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	Human driver and system	Human driver	Human driver	Some driving modes
2	Partial Automation	the *driving mode*-specific execution by one or more driver assistance systems of both steering and acceleration/ deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	**System**	Human driver	Human driver	Some driving modes
***Automated driving system* ("system") monitors the driving environment**						
3	Conditional Automation	the *driving mode*-specific performance by an *automated driving system* of all aspects of the dynamic driving task with the expectation that the *human driver* will respond appropriately to a *request to intervene*	System	**System**	Human driver	Some driving modes
4	High Automation	the *driving mode*-specific performance by an automated driving system of all aspects of the *dynamic driving task*, even if a *human driver* does not respond appropriately to a *request to intervene*	System	System	**System**	Some driving modes
5	Full Automation	the full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*	System	System	System	**All driving modes**

Figure 2: Illustrations are describing different levels of automation. SAE taxonomies focus on four aspects of driving (columns) to distinguish levels of automation. Levels 0 to 2 describe the levels where the driver is the principal acting agent of driving. Levels 3 to 5 describe the levels where the automated systems are the principal acting agent in driving.

For example, Level 3 of automation is identified as the Conditional Automated Driving System. It will provide full control of all safety-critical functions with occasional cases for the driver to intervene (SAE International, 2018). Vehicles at this level of automation enable the driver to cede full control of all safety-critical functions under specific traffic or environmental conditions and in those conditions to rely heavily on the vehicle to monitor for changes in those conditions requiring transition back to driver control. The driver is expected to be available for occasional control, but with a sufficiently comfortable transition time. Therefore, Level 3 systems would benefit from understanding driver's mental state and ability to handle an anticipated incident to plan the handover process by estimating driver's ability to respond to critical incidents.

While full automation is in action, drivers may direct their attention away from driving to engage in secondary tasks as seen in some studies (Merat et al., 2012). Nevertheless, Level-3 systems are still limited and will require the driver to re-engage within a predefined period of time in driving to handle a critical latent hazard (SAE International, 2018). This could be due to either sensory or decision-making limitations (Zeeb et al., 2015). Another scenario for the takeover is when a driver decides to switch the vehicle back to the manual system to enjoy driving. Though, drivers may not be ready for the transition due to their lack of situational awareness (Wright et al., 2016b). This transition of control (from automated to manual) is the critical bottleneck in the human-machine interaction. In the following section, the limitations and challenges of the transition of the control process are considered.

1.4 The Transition to Manual Driving

The term 'transition' has been widely used in the literature along with other terms such as 'handover' to refer to the transition process from automated to manual driving. A clear definition to the term '*transition*' is: "*The process and period of transferring responsibility of, and control over, some or all aspects of a driving task, between a human driver and an automated driving system*", (Louw, 2017). This definition aligns with Merat *et al.*, (2014) as the transfer of responsibility, or as the period taken to change from one vehicle state control to another (Flemisch et al., 2012).

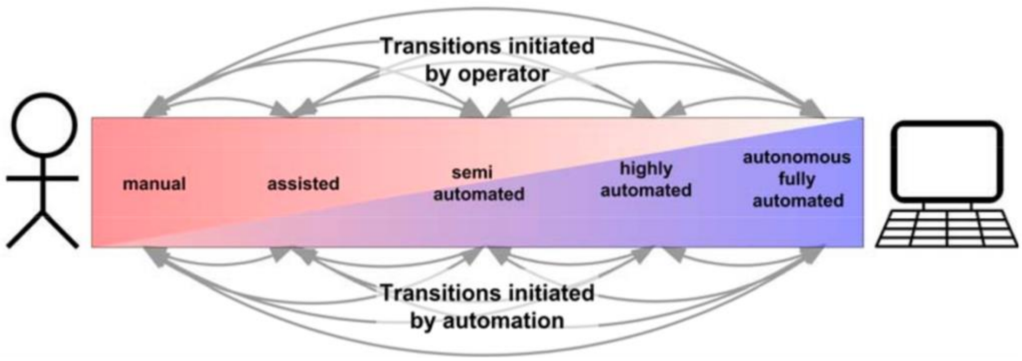


Figure 3: An illustration of all transitions possible between the human driver and the automated system (Flemisch et al., 2008).

Flemisch et al., (2008) introduced a few principles to shape the definition of the transition process. The first principle defining the transition is the flow of control between the driver and the different levels of automation. For example, the system could go from Level-3 automation to Level-1. This principle defines the transition based on the transfer of control between the human and the machine. Figure 3 shows all possible transitions among the five levels of automation.

The second principle of Flemisch et al., (2008) shapes the transition's definition by who has the control at the start of the transition and who is the recipient of control at the end of the transition. As illustrated in Figure 3, the transition may be given from one level of automated to another without necessarily involving the human driver in the process and for example, switching from Level-3 to Level-4.

The third principle of transition as defined by Flemisch et al., (2008) is the initiator of the transition. The initiation is triggered by one of the two agents involved in the 'transition', i.e., the automated system or the human driver. There's a distinctive difference between the two transitions because the human driver's readiness may be inconsistent. For example, a transition initiated by a human driver entails that the driver may be aware of the system limitation or is in full situational awareness (Larsson et al., 2014). However, a transition initiated by the automated system is often referred to as a 'mandatory transition' and is usually initiated due to a system limitation that may lead to a critical incident (Saffarian et al., 2012). This limitation could be due to the lack of driver's situational awareness or their mental capacity to handle a system limitation or critical hazard (Endsley and Kiris, 1995; Merat et al., 2012).

Several studies have discussed the issues of the 'mandatory transition', (De Winter et al., 2014). First known study to discuss this issue was Endsley and Kiris's, (1995). The main recommendations of Endsley and Kiris's, (1995) study was that the automated system has to put the driver's mental state into consideration before initiating the transition of control. In the past decade, several studies identified a substantial number of variables influencing the driver's performance during the transition phase (De Winter et al., 2014). A well-cited study argued that the driver's situational awareness is the main challenge to the performance of drivers (Merat et al., 2012).

Moreover, (Louw, 2017) recommended that "*Should the system be equipped with a driver monitoring system, the decision to relinquish driving control would have to be* *based on some empirical data of drivers' capacity and behaviour, in such conditions. For example, if the pattern of drivers' visual attention in the lead up to a transition shows that they were completely disengaged from the driving task, then a take-over-request could be delayed until drivers' attention is back on the driving task. Otherwise, the vehicle may initiate a minimum risk manoeuvre, bringing the vehicle to a safe position on the road. At present, this data does not exist. Therefore, it is further motivation to investigate drivers' capabilities and limitations in research*". This highlights one of the research gaps in the literature which is further explored in the literature review chapter.

Both situation awareness and mental workload levels are attainable through several physiological data measures. De Winter *et al.*, (2014) performed a comprehensive review of the partially controlled and highly automated vehicles. Their results identified physiological measures such as heart rate, eye movements, blinking and other features as reliable indicators to driver's mental workload and situational awareness. A drawback to their review is that it is out of date since the field has accelerated since 2014 in addition to their focus on partially controlled studies because there were not enough studies about highly automated driving. A comprehensive literature review is required to validate their findings and update them with the latest research in the past years.

1.5 Aim

The main aim of the study is to assess the correlation between the drivers' physiological behaviour and the quality of their performance during a transition from highly automated driving to manual driving. The following questions formed the basis of the conducted research:

1. What physiological data that could be collected in a highly automated driving environment to provide an assessment of the driver's response to critical incidents?
2. How does automation affect the collected physiological patterns of drivers in highly automated driving scenarios?
3. How do secondary tasks reflect on the driver's physiological patterns pre-transition, during transition and post-transition period?
4. What are the suitable driver performance measures to assess their responses during the 'mandatory' transition period?
5. What's the relationship between physiological data and driver's performance during the transition?
6. What features could be extracted from physiological data that could support the predictability of driver's performance?
7. How could physiological data be used to assess the predictability of the driver's performance?

1.6 Objectives

The research aim is achieved through the following objectives:

1. To conduct a critical review of existing literature to study different approaches for assessing drivers' physiological behaviour and its effect of their performance during the 'transition' phase. The objective is broken into the following stages:

- To review highly automated driving studies to identify the research gap and assess the different literature approaches used for data collection, analysis, and their data analysis methods
- To review different factors affecting driver's response quality in highly automated driving studies.

2. To design a driving scenario to assess driver's response quality during both manual and highly automated driving. The driving scenario involves the transition from highly automated to manual driving to understand driver's physiological behaviour pre- during and post-transition.

    - To write all the necessary code and acquire essential equipment to conduct the study designed on objective 2. The demographics of those participants are determined based on the outcomes of the literature review produced on objective 1 and on their suitability for the case study.

3. To conduct the study, produced on objective 3, on recruited participants.
4. To define an evaluation framework that assesses the efficiency of the prediction model produced by objective 6.
5. To assess the correlation between physiological patterns and driver's performance.
6. To develop a system that will take data collected from objective 4 to determine the outcome of the takeover done by the driver. The system includes a model capable of predicting drivers' response concerning time and quality before a takeover request. The efficiency of the model is further assessed using the evaluation framework developed on objective 5.

1.7 Thesis Overview

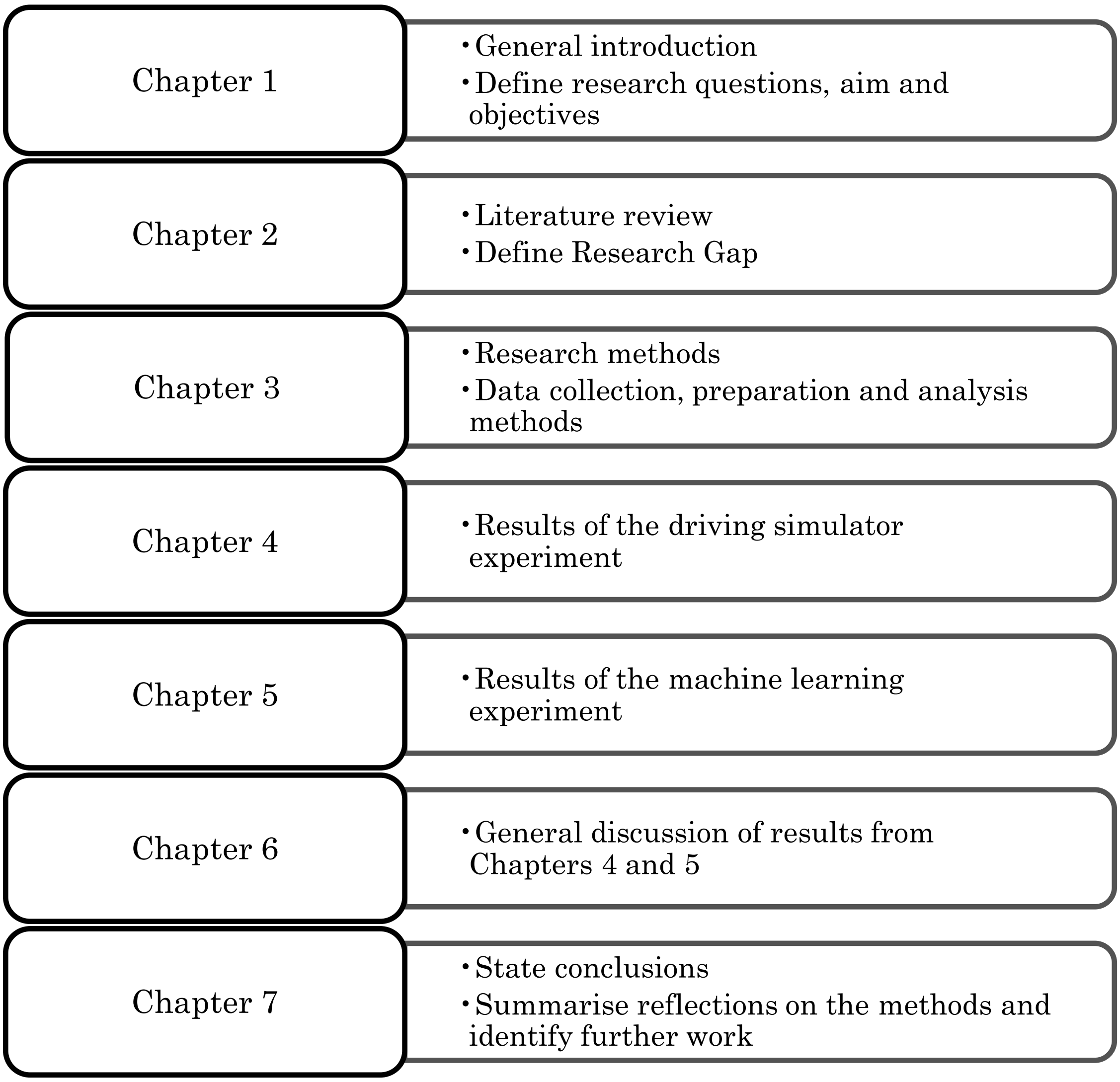


Figure 4: Thesis Structure.

- **Chapter two** reports a comprehensive literature review of automated driving to establish a clear understanding of the transition from automated to manual driving. Other studies from the manual driving field are reviewed to project missing research points in the highly automated driving field. Detection methods of driver's inattention are surveyed. In addition, physiological behaviour of drivers in highly automated driving (HAD) is surveyed and summarised. Finally, the research gap is defined.

- **Chapter three** details the research methods used in this study. Research philosophy, approach, data collection and analysis are explained and justified.
- **Chapter four** presents the results of the driving simulator study during the automation and the transition. Physiological behaviour of drivers is studied and correlated to their performance during the transition period.
- **Chapter five** presents the analysis of a machine learning-based approach in predicting the driver's performance. Results of regression and classification methods are explained.
- **Chapter six** presents the analysis of the results in Chapters four and five, correlated to each other and to other studies in the field. The physiological behaviour of drivers is analysed and correlated to their subjective and objective measures.
- **Chapter seven** summarises the results and discussions reported in Chapter four, five and six. A reflection on the methodology, data collection and analysis are reported. Finally, further work suggestions are provided.

2

CHAPTER

Literature Review: Highly Automated Driving Studies

Research studies proved that drivers tend to get ‘out-of-the-loop’ when using the level-3 automated driving systems (e.g. Endsley and Kiris, 1995; Merat *et al.*, 2012; Gold, Damböck, *et al.*, 2013; Körber and Bengler, 2014; Merat and de Waard, 2014; Radlmayr *et al.*, 2014; C. Gold *et al.*, 2016; Körber *et al.*, 2016; Louw *et al.*, 2016). The out-of-the-loop phenomenon is defined as a complete distraction from the driving environment, i.e., losing situational awareness. Loss of the situational awareness causes drivers to make poor decisions, especially when handling a critical handover caused by a system failure or limitation (Brookhuis and de Waard, 2001). Therefore, several studies focused on understanding what determines the handover time and the assessment of a driver's decision-making during the handover process.

In this chapter, issues of Level-3 automated systems are critically reviewed with a specific focus on the ones arising during the handover process. Generally, studies reviewed here focussed on the handover phase from an automated level-3 to manual driving (Level- 1 and 2). Each study focussed on an individual condition and how it impacted performance. Examples of such conditions are traffic density, secondary task, weather, and driver background measures (Merat and de Waard, 2014).

2.1 Differences among Transition, Handover and Takeover

Several studies tend to use handover, 'transition of control' and takeover terms interchangeably. A critical distinction between handover and takeover is identified in Merat's and de Waard, (2014) that defines handover as the process of transferring control from the vehicle to the driver starting with the takeover request until the driver has full control of the vehicle. The 'transition of control' has a similar definition; however, it refers to the transition between the human driver and the automated driving system in both directions (Louw, 2017).

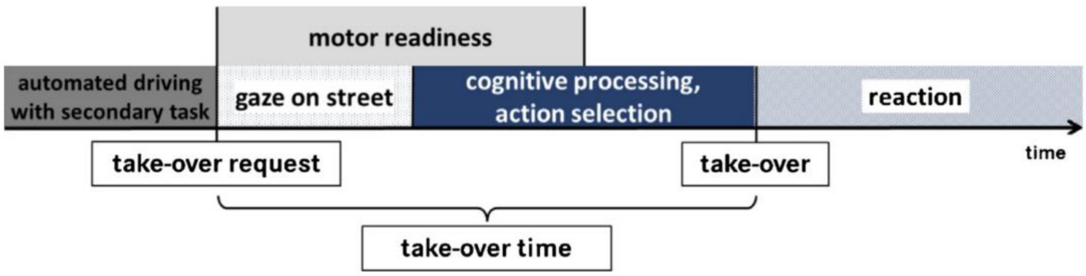


Figure 5: Model of the handover process showing when take-over time starts and ends (Zeeb et al., 2015)

Moreover, Merat and de Waard, (2014) defined the takeover as the specific time when the driver is in control of the vehicle. The main difference is that handover is a generalised term that covers the takeover request, transfer of control, and the time taken until the driver has gained full situational awareness. Conversely, the 'takeover' is a specific term that refers to the time a driver takes to regain control of the vehicle only.

In Figure 5, handover and takeover are illustrated on a timeline indicating the start and end of the takeover process during the overall handover process. As seen in Figure 5, the takeover is a subset of the handover process starting at the takeover request and ending at the time in which the drivers gain full control of the vehicle. To summarise, the study uses the term ‘transition of control’ to refer to the process of transferring some or all aspects of control between the human driver and the vehicle in both directions. The handover is the process of the automated driving system delegating some or all the driving task to the human driver.

2.2 Understanding Highly Automated Driving Studies

Studies in the literature utilised driving simulator technology to test simulated events or incidents safely on human participants. Each study has a driving scenario that ranged between 15 to 90 minutes in urban and suburban areas (Gold et al., 2016, 2013a; Jamson et al., 2013; Louw et al., 2016; Merat et al., 2012; Merat and de Waard, 2014). Short scenarios include a practice driving session before the study is taken (Merat et al., 2012; Neubauer et al., 2012). This is due to an observed learning curve to both the simulator and the automated systems (Körber and Bengler, 2014; Larsson et al., 2014; Wright et al., 2016a). Most studies recruited participants based on specific criteria such as years of driving experience, high annual mileage driven, and experience with Adaptive Cruise Control systems to satisfy the aim of their studies (Larsson et al., 2014; Wright et al., 2016a).

Experimenters in the reviewed studies in this chapter usually asked drivers to perform both manual and automated driving to allow for comparing different driving behaviours. One or more incidents or hazards are introduced in the middle or near the end of the study, to study driver’s ability to handle the critical incident. The experimental design varies environmental parameters to study their influence on driver’s performance. The parameters manipulated in those experiments could be weather based such as light or heavy fog (Louw et al., 2016), automated system design based such as time budget (Gold et al., 2013a), or driver based variables such as the influence of distraction (Zeeb et al., 2016). During some experiments, drivers are asked to perform a non-driving related (NDR) task to allow for a quantifiable distraction during the automated period of the scenario.

The NDR task could be a cognitive performance task such as the n-back task or Twenty-Questions Task (Radlmayr et al., 2014), a visually demanding task such as Surrogate Reference Task (Gold et al., 2013a), a simulated phone conversation (Körber et al., 2016) or naturalistic tasks such as watching a video, reading news or writing an email (Zeeb et al., 2016). Based on the collected data, experimenters evaluate how drivers responded to these critical incidents based on several factors affecting their decision making performance (Larsson et al., 2014; Merat et al., 2012; Radlmayr et al., 2014), age influence (Körber et al., 2016), or variables correlating with their cognitive workload load such a blink frequency (Merat et al., 2012), NDR tasks performance (Gold et al., 2016), and gaze behaviour (Louw et al., 2016; Zeeb et al., 2015).

Studies also look for other factors such as mental workload (Zeeb et al., 2016), situational awareness and driver opinion towards the system after conducting the experiment (Merat and de Waard, 2014) to have a broad understanding of drivers' behaviour and find the reasoning behind their performance. Finally, some studies investigated the influence of external conditions on drivers' performance during the handover process such as different traffic density levels (Körber et al., 2016; Radlmayr et al., 2014) and simulated weather conditions (Louw et al., 2016).

The study performed by (Merat et al., 2012) is a good example of the standard methodology used by most studies in the field. The main objective of the study was to compare the differences in the mental workload of drivers between Level-1 and Level-3 driving. The driver's performance was evaluated based on their blinking patterns. The blinking pattern performance measure was chosen due to its ease of use and non-intrusiveness nature in detecting the driver's workload, fatigue, and stress (Neumann and Lipp, 2002). The blinking frequency was collected using a FaceLab eye tracker (faceLAB, 2016). The scenario of the study was to ask participants to drive on the same road both in Level-1 and in Level-3 modes. In Level-3 mode, some drivers were asked to perform an NDR task to elevate their cognitive workload. The NDR task used in the study was the Twenty-Questions Task (TQT); explained in detail in section 2.3.1.2. Drivers were also expected to drive manually on the same road to compare their blinking patterns in Level-1 and Level-3 modes. The study recruited 50 participants to perform on a driving simulator. Each participant had a 45 minutes' practice session to get familiar with the simulator on manual and Level-3 automation levels. In between the two sessions, drivers had a break in between to alleviate fatigue potentially caused by the long sessions.

Half of the participants of Merat *et al.*, (2012) started the Level-1 session first while the other half started the Level-3 session first to minimise the order effect. During the Level-3 mode, participants were asked to perform the TQT, i.e., guess a specific object by asking the experimenter a maximum of twenty yes-no questions. This TQT lasted for 3 minutes and was performed twice. One of those times, the TQT was followed by a critical incident that required driver's intervention. Collected data showed that blinking frequency patterns were much lower during the high workload periods. Also, the blinking frequency was more consistent during Level-1 in comparison to Level-3 mode which had a much higher inconsistency in blinking frequency. Therefore, the study concluded that driver's performances had no significant differences in both modes; however, when a secondary task is introduced, the performance of drivers is highly degraded during the takeover process.

To conclude, this section introduced an overall summary of the literature of highly automated driving. The key questions faced by the researchers in the field were briefly discussed and their findings were reviewed briefly. The structure of experiments in the highly automated driving field was introduced in detail based on Merat's *et al.*, (2012) study. Several factors affecting driver's performance were introduced briefly. More details will be provided in section 2.3.

2.3 What Affects Driver's Response?

In the following section, several studies that investigated different aspects affecting the takeover time and performance are critically reviewed. The reviewed studies are divided into six categories. The first category discusses the time budget given to drivers to responds after the takeover request and its effect on the quality of driver's responses. The second category is named 'driver distraction', and it includes all studies inducing distraction on drivers using NDR tasks. It is grouped by the NDR tasks performed prior to the Takeover Request (TOR) which could influence the handover process. In this category, several NDR tasks were introduced to affect the driver's cognitive and visual workload. The third category is named 'stress and fatigue', and it includes all studies inducing fatigue or stress on participants in the highly automated driving environment. The fourth and fifth categories include the manipulation of traffic situation and road conditions. The reviewed studies of the two categories investigated the influence of road and traffic conditions on the driver's performance during the handover process. The sixth category is named human machine interfaces, and it discusses the impact of the communication design between the automated system and the driver. Finally, the seventh category is named 'driver background' which include studies focusing on driver's global factors such as age, driving experience in addition to their individual differences in handling acceleration and braking. In the following sections, each one of them is critically reviewed. A quick summary of the definitions is laid out on Table 1. Then, highlights of the main studies reviewed are summarised in Table 2.

Table 1: Summary of the six categories of factors affecting driver's responses in highly automated driving studies.

Category name	**Summary**
Time budget	Time budget is the time between the takeover request and the time a critical incident occurs if the driver doesn't apply any changes to the vehicle's lateral or longitudinal speed.
Driver distraction	The driver distraction category is defined through the type of tasks the driver is distracted with before the takeover request.
Fatigue and stress	The stress level and fatigue of drivers influence their responses during the takeover process. Stress or fatigue could be caused by driving- or externally-imposed (pre-driving condition).
Traffic situation	The traffic situation category is concerned with any factors influenced by nearby traffic, whether it's traffic density or emergency vehicles such as police cars, etc.
Road conditions	The road conditions category is concerned with any factors influenced by the weather, road shape or visibility,
Human machine interfaces	The human machine interfaces category is concerned with the factors affected by the design of the vehicle's cockpit and its communication channels with the driver such as communicating uncertainties, etc.
Driver's background measure	The driver's background measure are the driver's personal factors such as age, driving experience and personal driving style.

2.3.1 Time Budget

Many studies varied the time budget given to drivers to handle the takeover situation in order to understand the driver's ability to gain full control after a takeover request (Eriksson and Stanton, 2017). The time budget or the time-to-collision is the time between the takeover request and the time the vehicle collides with another vehicle or object if the driver doesn't interfere. Most studies focused on near-crash scenarios on a motorway while varying different variables such as traffic density, drivers experience and other factors (Eriksson and Stanton, 2017). The first known study to vary the time budget was Damböck *et al.*, (2012) who evaluated 4, 5, 6 and 8 second time budgets. The study reported a significant crashing rate in groups that were in the range of 4-6 seconds in comparison to the 8 seconds group. More studies investigated the time budget in depth.

The driver's response time differs significantly among studies (e.g., Gold et al., 2016, 2013; Zeeb et al., 2015). Literature showed that response time is influenced by NDR tasks (Merat et al., 2012), drivers' background measures (Körber et al., 2016), driving experience (Wright et al., 2016b), and road conditions (Gold et al., 2016; Radlmayr et al., 2014). Among those, NDR tasks are the most dynamic factors influencing response time. Several studies reported seven (Gold et al., 2013a), eight (Wandtner et al., 2018), ten (Melcher et al., 2015) and 12 seconds (Zeeb et al., 2015) as a safe time budget to be given to drivers to respond.

Intriguingly, studies reported that drivers take a longer time to respond if they were given a more extended time budget (Gold et al., 2013a). An explanation to this could be due to drivers investing time in restoring situational awareness before taking an action; hence, they had a safer response as measured by the objective measures in comparison to the group who were given a shorter time to respond (Gold et al., 2013a; Radlmayr et al., 2014; Zeeb et al., 2016).

Moreover, studies recommended that shorter time budgets increase the probability of crashing. For example, van den Beukel and van der Voort, (2013) reported that shorter time periods increased the crashing probabilities significantly. Their study found that 47% of drivers with a time budget of 1.8 seconds were unable to avoid a collision in comparison to 12.5% of drivers with 2.8 seconds. The findings of van den Beukel and van der Voort, (2013) aligns with Zeeb, Buchner and Schrauf, (2015) where they reported 45% and 15% of drivers crashing when given a time budget of

4.9 and 6.6 seconds respectively. The findings of the two studies suggest that human drivers perform very poorly in a time-restricted take-over scenario.

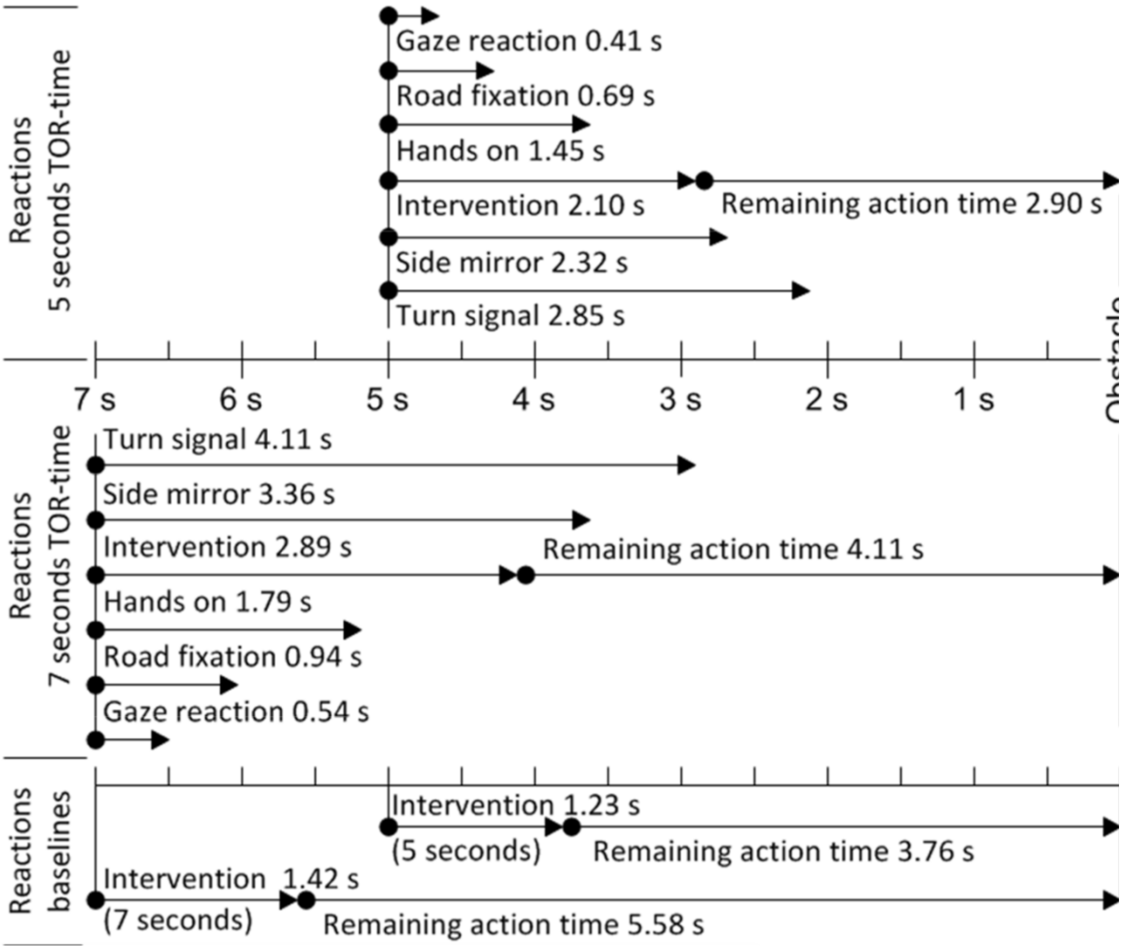


Figure 6: Illustration of how drivers allocated their time budget to respond to a critical hazard Gold, Damböck, et al., (2013).

The findings of Zeeb, Buchner and Schrauf, (2015) and van den Beukel and van der Voort, (2013) attempted to extend and challenge previous findings of Gold, Damböck, *et al.*, (2013). The study conducted by Gold, Damböck, *et al.*, (2013) examined 5 or 7 seconds time budget until a collision occurs with a broken vehicle in the ego-lane. Their study's baseline was a group of drivers performing the same task in manual driving. Gold, Damböck, *et al.*, (2013) crafted a detailed understanding of driver's behavioural response to a critical hazard with significant details to the spending of their time budget as illustrated in Figure 6. The study found out that drivers were given 5 seconds to respond reacted faster to the hazard, but they were much more likely to skip critical safety checks such as glancing at the rear and side mirrors before their lane change manoeuvre. The study concluded that a 7 seconds budget is adequate for drivers to respond to critical hazards in highly automated driving. Subsequently, the results from Gold, Damböck, *et al.*, (2013) have been adopted in recommendations by the NHTSA as a standard for manufacturers to design their automated system's transitions (Campbell et al., 2018).

2.3.2 Driver Distraction

Driver distraction is the process of disconnecting from driving and is categorised as visual, cognitive (Lee et al., 2001) or manual (Craye and Karray, 2015) distraction. Visual distraction is identified as the loss of visual concentration on the road for a quantified period. Visual distraction is quantified in driving studies as the eyes-off-road time such as watching a video or reading a document (Craye and Karray, 2015). Visual distraction is caused by either cognitive or visual causes. Visual causes are a result of onboard presence and multimedia devices and salient visual notifications of driving-related activities. For example, an entertainment system, a petrol level warning light or over-speed notification on the dashboard might cause spontaneous off-road glances (Haigney and Westerman, 2001).

Cognitive distraction is another factor which is identified as the insufficient concentration of a driver on a critical driving task. It concerns cognitive processes and has been described in a few studies as the mind-off-road (Liang and Lee, 2010; Victor, 2005). For example, it may occur when the driver is talking to other passengers, on the phone or during texting (Craye and Karray, 2015). The symptoms of cognitive distraction are less apparent and harder to detect or quantify in comparison to visual distraction. Therefore, they require more sophisticated techniques and long-term detection of the driver's patterns (Recarte and Nunes, 2003; Zhang et al., 2004).

Manual distraction involves hand activities such as holding a cup or a phone. Even when eyes are still on-road, the driver's response time is longer in comparison to a situation without manual distraction. This delay is thought to be caused by the additional mental and manual effort to get rid of the manual distraction before engaging in the driving task (Craye and Karray, 2015).

With level-3 automated driving systems, drivers may engage in distracting tasks that will acquire their visual and cognitive attention (Cantin et al., 2009; Hancock et al., 2009). The distraction of drivers raises several safety concerns when the automated systems signal a takeover request (TOR) to drivers to handle a system limitation or a latent hazard. To understand the consequences, studies assessed how different tasks might affect driver's situational awareness before the takeover request, their readiness for the transfer of control, and the restoration of situational awareness starting from the TOR until the end of handover process (Merat and de Waard, 2014). Such situational awareness involves perception, comprehension of the conditions and projection of the latent hazard (Endsley, 1995).

During a takeover, drivers responses to critical incidents were comparable to their driving behaviour in manual driving when they are not performing any NDR tasks (Merat et al., 2012). Once an NDR task is introduced, drivers' responses degraded significantly (Gold et al., 2016, 2013b; Healey et al., 2012; Merat et al., 2012; Merat and de Waard, 2014; Radlmayr et al., 2014).

At an abstract level, drivers are usually distracted by a task that possesses most of their cognitive, visual or manual attention. The NDR tasks make it more difficult for a driver to regain the situational awareness due to the high workload demanded by the secondary task (Baumann et al., 2007; Blanco et al., 2006; Kass et al., 2007).

Similarly, response quality is impacted by the NDR tasks due to their visual or visuo-cognitive distraction from the driving environment (Zeeb et al., 2016). To respond to a take-over situation, drivers require a cognitive processing time to restore situational awareness and respond accordingly (Endsley, 1995; Endsley et al., 1997).

In the automated driving literature, studies investigated several NDR tasks that induce cognitive workload such as n-back task (Radlmayr et al., 2014), Twenty Questions Task (Merat et al., 2012), visuo-cognitive tasks such as reading news (Zeeb et al., 2016), internet search (Zeeb et al., 2016), vehicle's multimedia systems (Zeeb et al., 2015) and IQ questions (Louw et al., 2016).

Gold et al., (2015) reported a significant decrease in performance for tasks including manual versus cognitive workload. Such results concur with the findings of Petermann-Stock et al., (2013) that the worst performance decrements were caused by quizzes requiring a combined visual, cognitive and manual workload in comparison to quizzes requiring one or two of those workloads. Their findings contradict Gold et al.'s, (2015b) findings that reported cognitive and manual tasks had the same detrimental effect. The contradiction provides the necessary motivation to study the physiological differences caused by cognitive and visuo-cognitive tasks in order to provide a better understanding of the differences among those studies. As mentioned earlier, critically reviewed studies in this section used several techniques to simulate visual and mental workload with

NDR tasks. The adopted NDR tasks are explained and discussed in the following sections, with selections that include cognitive and visuo-cognitive tasks.

2.3.2.1 N-back Recall Task

The N-back Task is a continuous performance task used to assess and challenge the capacity of the working memory (Bruce Mehler et al., 2011; Kirchner, 1958). "The auditory attention and memory components of the task draw on many of the same cognitive resources utilised when engaging in an externally paced task such as responding to a cell phone call or interacting with an in-vehicle device that uses audible prompts or control commands. Similarly, it draws on cognitive resources that are utilised for less structured interactions such as attending to and maintaining a conversation with a passenger", (Mehler et al., 2011).

The n-back task is performed by asking the participant to listen to a set of digits and then asked to say aloud the nth digit from last digit (Mehler et al., 2011). For example, Louw et al. (2016) used the 1-back task to induce a cognitive workload on drivers in a typical driving environment. Participants were asked to repeat aloud the last single digit number they heard in a series of numbers. 2-Back Task was used by Radlmayr et al. (2014) to compare the visual (caused by the Surrogate Reference Task) and cognitive distractions caused by the 2-back task to assess the quality of driver's takeover. The Radlmayr et al. (2014) study has shown the same influence of both tasks on dense traffic scenarios. In summary, the N-Back task is considered in studies that analyse the cognitive workload effect on drivers without necessarily engaging them visually.

2.3.2.2 Twenty-Questions Task

The TQT is used to stimulate cognitive reasoning and creativity by asking participants a series of questions that require reasoning and memory recall to reach a specific answer or make an informed estimation (Walsorth, 1882). In driving studies, it is used to induce the driver's mental workload to stimulate cognitive distraction. For example, Gold et al. (2016) and Körber et al. (2016) used the TQT to make their participants guess an animal by asking the experimenter a series of polar (yes-no) questions; when the animal is guessed correctly, the participant is asked to guess a new one to continue the same level of mental workload.

A study by Strayer, Drews and Johnson, (2003) compared the mental workload of the TQT of drivers in highly automated driving environment and a manual driving group. They reported that the mental workload imposed by the TQT on HAD drivers was comparable to the mental workload imposed by manual driving. It also reduced situational awareness. As a result, it prolonged driver's reaction time because the driver required an additional time to gain a full understanding of their surroundings (Strayer et al., 2006). The study of Gold et al. (2016) reassured that the TQT has an adverse effect when merged with high traffic density. The Gold et al. (2016) study concluded that drivers who are distracted by the TQT take longer to react because they require a longer time to regain full situational awareness during the takeover manoeuvre. They also performed poorly with a higher number of collisions and near collisions in comparison to participants who were not distracted by the TQT task on the same experiment.

2.3.2.3 Simulated Hands-free Phone Call

Some studies merge the TQT with simulated hands-free phone calls (HFPC) to make it more naturalistic for drivers. The literature showed that HFPC has no significant effect on increasing chances of safety-critical incidents (Fitch et al., 2013); however, accompanying it with the TQT raises those chances (Heenan et al., 2014). Merat et al. (2012) used a simulated hands-free phone call with the TQT for verbally guessing questions. Results showed that the worst performance of the takeover process was when drivers were not allowed to interrupt the task during the handover process. The HFPC and N-Back Tasks were used on elderly groups to challenge their mental allocation, pausing, and resuming of driving and NDR task ((Körber et al., 2016; Radlmayr et al., 2014). Radlmayr *et al.*, (2014) used the N-Back task and reported that age is a strong influence of driver's performance while Körber *et al.*, (2016) used the TQT and reported that age had no influence. These conflicting outcomes are probably influenced by the researcher's choices of the NDR tasks. It's likely that the N-Back Task posed a stronger cognitive demand in comparison to the simulated phone call. Therefore, a study in this field might consider the consequences of the choice of their NDR tasks. More details about Körber *et al.*, (2016) and Radlmayr *et al.*, (2014) are provided in 2.3.7.1 section.

2.3.2.4 Surrogate Reference Task

The Surrogate Reference Task is a visually demanding task that involves searching and responding; for example, participants are presented with a group of circles (typically on display) and are asked to identify the biggest one (Jamson and Merat, 2005). Gold, Damböck, Lorenz and Bengler (2013) used SuRT as a secondary task to visually distracting drivers to determine their response time when the automated system issues a TOR. The task was enhanced with a score graph to engage drivers in it. Based on Radlmayr et al. (2014), SuRT has a similar effect to n-back task in heavy traffic scenarios.

2.3.2.5 Naturalistic Tasks

Naturalistic tasks simulate the most common scenarios that may occur naturally during the automated driving; hence they were common among HAD studies. Some studies used naturalistic tasks to simulate distraction in driving scenarios. Reading (Wright et al., 2016b; Zeeb et al., 2016), writing an email, watching a video (Zeeb et al., 2016), Internet searching and texting (Zeeb et al., 2015) were examples used in the reviewed literature. Naturalistic tasks were explored in manual driving studies. For example, the effects of eating, drinking and grooming were studied (Sayer et al., 2005). The aforementioned tasks are yet to be explored in the highly automated driving studies.

A major drawback to naturalistic tasks is their inconsistent level of distraction on participants. For example, in-vehicle entertainment (sweets, hand-held games, magazines, and films) was used in Jamson et al. (2013) to encourage drivers to get distracted. They reported that experienced drivers spent less time distracted by the entertainment system in heavy fog in comparison to light fog environment which was due to a sense of responsibility towards their safety. However, experienced drivers were more distracted generally in comparison to their state in manual driving.

Some other studies explored IQ-based questions. Louw et al. (2016) challenged drivers cognitively with a quiz based on IQ questions to assess shape matching, general knowledge, and moderate mathematical problems. Drivers were going through a simulated fog environment, and the quiz was used to induce their cognitive workload. While this is a relatively new approach, the paper has provided neither the questions nor the selection criteria used to pick out those questions. Therefore, replicating the study would be very difficult.

2.3.2.6 Tactile Detection Response Task

Tactile Detection Response Task (TDRT) is another recently developed response task to assess driver's response time. A vibration signals drivers to touch a tactile surface to acknowledge the signal within a time limit (Young et al., 2013). The time taken from the signal to the touch gives a cue on the alertness of the driver and their response time.

HAD studies have not used TDRT in their studies. However, some performed a similar approach to detect the driver's motor readiness. For example, participants of Zeeb, Buchner and Schrauf, (2016) pressed a button in the steering wheel at the start of a takeover to gain back full control of the vehicle; hence, measuring the motor readiness time of drivers. In practice, Zeeb, Buchner and Schrauf, (2016) used the TDRT approach implicitly.

2.3.2.7 Summary of Driver Distractions

In summary, HAD studies focused on using TQT as a secondary task especially in scenarios that required adding a cognitive workload on drivers without imposing any visual distraction. SuRT was used in scenarios that required visual distractions while the n-back task was used to induce a strong cognitive distraction. Naturalistic tasks were used but not as frequently because of their inconsistency in inducing visual or cognitive workload in comparison to TQT and SuRT.

To satisfy the main aim of this study, driver's behaviour under the influence of visual and cognitive distractions separately should be collected and analysed to train and evaluate the performance of the prediction model in both scenarios. To conclude, each secondary task affects the driver's attention cognitively and visually, and the selection should be based on the scenario and hypothesis of the study, please check section 3.3.5 for more details on the study's approach in selecting the suitable NDR tasks for the experiment.

2.3.3 Fatigue and Stress

Fatigue is a result of physical, physiological, or psychological causes as it is correlated to drowsiness. Its symptoms could be drowsiness and frequent nodding, and it is caused by either external factors such as the driver's preconditional state before driving or due to long periods of driving (Liang and Lee, 2014).

Fatigue and stress are key factors impairing driver's performance in automated driving (Morgan et al., 2016) and such automation may exacerbate drivers' fatigue which in turn impairs their safety-critical performance to handle immediate hazards (Saxby et al., 2013). Rauch et al. (2009) highlighted the importance of a drowsiness and fatigue detection system in HAD vehicles. The study also identified fatigue as a hindering factor against the driver's ability to get back-in-the-loop during the handover process.

To detect fatigue, Percentage of Eyes Closed (PERCLOS) is considered an accurate measure of drowsiness and fatigue in driving situations. It measures the percentage of time when eyes are closed and based on a threshold; it detects fatigue (Wierwille et al., 1994). Jamson et al. (2013) used PERCLOS to prove that driver's fatigue rises from 1.8% in manual driving to 3.8% in automated driving. The observed increase indicated a decline in driver's arousal caused by the automation. However, the same phenomenon was not observed in heavy traffic conditions during the experiment. Drivers were expected to drive for 35-minutes following a 40-minute practice in a simulator. Even though long experiments allow for more data, drivers seem to reach a drowsiness level after 90-minutes of a HAD study (Alford, 2009; Morgan et al., 2016); it is not known whether the drowsiness is caused by the driving routine or by the experiment effect itself.

Therefore, a road study should be done to understand how long driving in HAD vehicles affects drowsiness level in drivers. Moreover, the effect of fatigue during extended driving sessions beyond 90 minutes is still to be explored. Finally, the effect of sleepiness and drowsiness should also be investigated during the handover process for both short and long driving sessions.

2.3.4 Traffic Situation

Traffic density has shown a strong influence on takeover time and performance. Gold et al. (2016) found out that traffic state has a significant impact on takeover performance during an emergency handover. This aligns with findings of Radlmayr *et al.*, (2014). The study of Gold et al. (2016) explained that the traffic density has a 'ceiling effect' in which the performance degradation reaches a peak when the density is around 15 cars per kilometre and remains the same as the number grows. The degradation could be explained by the driver's extended visual scanning and longer decision-making process. The study suggested further exploration of the relationship between traffic density (between 1 and 15 cars per kilometre) and handover performance. Gold et al. (2016) conducted the study on 72 participants from a younger [Mean (M) = 23.3, Standard Deviation (SD) = 2.6] and older [M = 66.7, SD = 4.56] and suggests that age was not a factor in the handling of traffic density. The study has not considered middle age participants [M=35-40] who may perform differently from the two age groups that were used.

2.3.5 Road Conditions

Road conditions were found to be another effective factor during the handover process. A study found a strong negative correlation between road visibility (i.e., fog heaviness) and the number of crashes. Those who performed a takeover in heavy fog were 46% more likely to crash in comparison to 33% in no fog scenario in the same experimental design. During the transfer of control in a heavy fog environment, erratic eye movements were identified in those who were more likely to crash in comparison to those with smoother eye pattern movement who were less likely to crash (Louw et al., 2016). This is because those drivers were able to identify the potential hazard earlier and therefore avoid it. However, it could be due to two reasons as stated by Louw *et al.*, (2016). First, it could be an increase in cognitive demand on the participants which was caused by an ‘automation surprise’ (Hollnagel and Woods, 2005). Automation surprise is defined as an action made by the automated system and wasn’t expected by the user (Hollnagel and Woods, 2005). The second reason could be due to driver’s over-trusting the system to handle the hazard (Lee and See, 2009). The literature has not yet explored other road conditions such as heavy rain and storms.

2.3.6 Human Machine Interfaces

The literature has explored several approaches to identifying optimal communication channels between drivers and automated driving systems (Helldin et al., 2013; Kunze et al., 2018a; Louw, 2017; Zhang et al., 2018). Studies also explored communicating the uncertainties of the system through visual cues (Kunze et al., 2018a) and light feedback (Kunze et al., 2018b) to keep drivers in the loop. This, in turn, improved trust and driver’s allocation of attention (Kunze et al., 2018a). Kunze *et al.*, (2018) introduced a new method to communicate the automated driving system’s uncertainties using a displayed graphics simulating heart rate frequency on the dashboard. The high frequencies represented a high uncertainty of the driving environment. The results of Kunze *et al.*, (2018) showed an improved response time, driver’s performance and minimum time to collision. Results aligned with the findings of similar studies (Beller et al., 2013a; Helldin et al., 2013). A main drawback to the Kunze *et al.*, (2018) is the use of a simple NDR task (SuRT) which may not have imposed a strong visuo-cognitive distraction on drivers. To summarise, the communication of the automated driving system’s uncertainty improved driver’s vigilance during the pre-TOR period which in turn improved their performance handling a critical incident. Human-machine interfaces were considered as out of the scope of the study; refer to section 1.5 for aim and objectives. Thus, advanced human-machine interfaces were not investigated.

2.3.7 Driver Background Measures

Several studies identified driver specific variables that had an influence on drivers' responses such as trust, mental capacity influenced by age (Körber et al., 2016) or driving experience (Larsson et al., 2014). In the next sections, the influence of driver's background variables is reviewed and discussed.

2.3.7.1 Age

Driving complexity, in addition to the complexity of the handover process, is challenging for older drivers (Anstey et al., 2005) and research regarding age is motivated by several reasons. Age causes a decline in cognitive processing (Salthouse, 2009), reduces adhered focus, lowers divided attention (Siu et al., 2008) and task switching (Kray et al., 2004). Due to those, ageing causes a slower response time when the participant is interrupted (Monk et al., 2004). Therefore, switching from Level-3 to Level-1 driving mode could be a challenge for older drivers. To explore this, several studies investigated ageing influence in both manual and handover process in automated driving.

Petermann-Stock et al. (2013) carried out a comparison study between a younger [25-35 years old] group and an older group [50-70 years old] during a high cognitive workload experiment. The study concluded a difference of up to 1200 milliseconds in reaction time. A limitation of the study is the lack of immediate hazard or condition required to demand a quick reaction time. Such limitations biased their results because participants were not prompted to take immediate action; hence, the observed difference between the two age groups.

Following up on this, Körber et al. (2016) carried out a more elaborate study with varying conditions during the handover process that overcame such limitations. Körber et al. (2016) also took into consideration the studies concluding that high traffic density causes a longer reaction time (Trick et al., 2010), especially for older drivers (Cantin et al., 2009; Horberry et al., 2006). Thus, Körber et al. (2016) used different variants of traffic density to assess whether it would influence participant’s takeover time in comparison to a younger group. The study used 72 participants in two groups; a younger group with an average age of 36 years old and an older group with an average age of 66.6. Participants were asked to perform a simulated hands-free cell phone conversation as engagement in a non-driving related task. Take-over time was defined as "the time between the TOR and the first conscious reaction by the driver, i.e. a change of 10% of the maximum brake pedal position or more than 2 degrees in steering wheel angle", (Gold, Damböck, Lorenz and Bengler, 2013). Results of Körber et al. (2016) showed no significant difference in the takeover time regardless of age or task. However, they reported that traffic density was the strongest influencer in extending the takeover time. This could be due to the time it took drivers to restore full situational awareness. For example, with more vehicles on the road, more time was needed by the drivers to perceive, comprehend and project the next step of each one of them. It was noticed that the number of accidents and takeover time decreased with every new takeover request for each participant indicating a learning curve for the system. 10% of both age groups got involved in an accident which indicated that age might not have a strong influence in the safety of the handover process.

In contrast, Radlmayr et al. (2014) found out that reaction time and quality of the study's participants (average of 35 years old) are highly degraded during the takeover process which contradicts with the findings of Körber et al. (2016). Radlmayr et al. (2014) used n-back task for half of their participants as a cognitively engaging continuous task before and during the takeover process (i.e., N-back task assesses if participants remember whether a current item is the same as the one presented n-items previously, explained in detail in section 2.3.1.1). However, Körber et al. (2016) used a simulated phone call employing Twenty Questions Task (TQT) that was interruptible during the takeover. TQT is a set of questions that stimulate deductive reasoning and creativity (Walsorth, 1882) and it is reviewed in section 2.3.1.2 . The conflict of their findings could be due to the different stimuli of the chosen secondary tasks and in allowing drivers to interrupt the task during the takeover process.

It is worth noting that a simulated phone call with the TQT task might not have imposed a strong cognitive challenge on older participants in comparison to the n-back task. A future study should compare effects of several cognitive tasks, natural tasks in addition to visually distracting tasks on handover process and how it may affect reaction time on several age groups to provide a more consistent comparison among different tasks in different age groups.

Strayer and Johnston (2001) conducted a similar experiment on manual driving context (Level 1 automation) and discovered a significant drop in driver's performance while engaging in a simulated phone call. Comparing findings of Körber et al. (2016) with Strayer and Johnston (2001), it shows different effects of the same secondary task on drivers while using Level 1 Automation systems on the one hand and Level 3 Automation systems on the other. Also, a study conducted by Guo et al. (2016) on drivers using Level-1 vehicles has indicated that "teenaged, young adult drivers [16-29] and senior drivers [65-98] are more adversely impacted by secondary task engagement than middle-aged drivers. Visual and manual distractions impact drivers of all ages, whereas cognitive distraction may have a larger impact on young drivers". Such findings align with the findings of level-3 automation studies; however, it raises a research gap in the field of Level-3 vehicles to compare the effect of long-term cognitive distraction on different age levels especially teenage and senior groups.

2.3.7.2 Driving Experience

Driving experience strengthens driver's comprehension to gain situational awareness faster and perform better decisions in comparison to novice drivers. For example, comparing two groups of experienced and inexperienced drivers showed a positive influence of experience on detection and anticipation of latent hazards during the handover process (Wright et al., 2016a). The proportion of glances spent on potential hazards during four transfer of control experiments were higher with experienced drivers than inexperienced ones which means experienced drivers could identify latent hazards faster and more efficiently. The study concluded that middle-aged drivers could anticipate 83% of latent hazards in comparison to 71% for inexperienced drivers. This could be explained by the presumption that experienced drivers could gain full situational awareness faster during the handover process.

Samuel et al. (2016) identified the adequate time for participants to gain situational awareness as eight seconds before a complete transfer of control. Their study recommended that eight seconds is the time the driver takes after the TOR to regain situational awareness and perform a driving decision. The study was conducted on a group of 18-22 years of age which could be arguably described as inexperienced drivers (Wright et al., 2016a). To expand on that, the study of Wright et al. (2016*b*) replicated Wright, Samuel, Borowsky, Zilberstein and Fisher, (2016) using a group of experienced drivers. Their results identified that 6 seconds were enough for middle age experienced drivers to gain full situational awareness before the transfer of control. Results of the study match with results of Louw *et al.*, (2016). The study of Wright et al. (2016*b*), Horswill and McKenna (2004) also suggested that latent hazard anticipation is a valid measurement of situational awareness assessment during the transfer of control . Latent hazard anticipation is defined as "the ability to detect and respond to potential threats that have not yet necessarily emerged on the forward roadway", (Wright et al., 2016a). It is also interesting to note here that Louw et al. (2016) recommends a future system to direct driver's visual attention immediately towards the hazard for a safer transfer of control.

Experience with advanced driving systems such as ADAS or ACC gives an advantage to drivers handling Level-3 automated driving systems. For example, Larsson et al. (2014) concluded that drivers with experience in ACC systems, level-2 automated systems, were 500ms faster in gaining back control than drivers who never used ACC systems. The study also suggests that ACC experienced drivers were proactive in taking control when they realised a system is behaving in an unsafe manner in comparison to inexperienced drivers who waited until they were instructed to take over control. This could be due to their prior knowledge of system limitations and their comfort in handling the transfer of control (Larsson et al., 2014). However, the study used only one driving situation to test their hypothesis in and failed to estimate when inexperienced drivers would gain such experience. The outcomes of Larsson et al. (2014) along with Wright et al. (2016*a*), Körber and Bengler (2014) suggest that an efficient HAD study should consider both drivers with and without ACC driving experience if possible in their recruitment to allow for unbiased results.

2.3.7.3 Individual Differences

Drivers have individual differences that shape their unique manoeuvring and controlling behaviours (e.g., brakes, wheel, and acceleration). Nilsson et al. (2015)

proposed the Driver Controllability Set, a statistical model that defines a subset of driver's behaviours, allowing the system to understand their manoeuvring style. The model is built automatically using statistical analysis of the driver's behaviours during manual driving. Using the model, a driver's transition from automatic to manual could be classified as safe. The study was performed on real-world data and proved to be suitable for real systems (Nilsson et al., 2015). While this might not provide a lead on takeover time, it is valid feedback for assessing the transition performance.

Table 2: Highlights of the main studies reviewed in the literature review of this study.

Study	Aim	Secondary Task/s	Participants	Experiment Preparations	Devices Used	Data Collected
(Gold et al., 2016)	Assess changes in timing and quality aspects during takeover when traffic density is manipulated, and the verbal TQT in a phone call is used to add a level of driver cognitive load.	TQT Questions asked to make the driver guess an animal.	19-79, median 24 and SD, 22.2. At least one year of experience	Some drivers were paid. A consent form was signed. A brief was given. The test drive was given.	Driver simulator. Head-mounted eye-tracker (Dikablis)	In three takeovers, the following data were collected. Timing aspect: Hands-on time, takeover time. Quality aspect: maximum longitudinal acceleration, maximum lateral acceleration, minimal time to collision. Workload assessment: horizontal gaze dispersion.
(Körber et al., 2016)	Influence of age on the take-over of vehicle control in highly automated driving	Hands- free phone call with TQT	72 participants. 80.6% males. Two age groups. The young group 19-28 and old group 60-79.	Questionnaire about age, gender and experience. Introduction and test drive.	BMW 6 series mock-up simulator	Vehicle position, acceleration, steering wheel, angle, and position of pedals at a frequency of 100 Hz.
(Radlmay r et al., 2014)	r How traffic situation affects takeover quality	Surrogate Reference Task and n-back task.	48 participants (38 males). Mean age is 33.5 and SD = 9.0	A demographic questionnaire, briefing and a test drive.	Dikablis Gaze Analyser and BMW high fidelity simulator	vehicle and situational parameters with video and audio recording in addition to gaze behaviour.

(Zeeb et al., 2015)	What determines take-over time	naturalistic tasks: internet search, search, texting	89 participants, (54 males), 20-72. Mean age 42	Demographic questionnaire, briefing and test drive.	Daimler AG dynamic simulator, Two cameras for eyes movements. Another	Gaze behaviour and body movement behaviour.
(Louw et al., 2016)	Visual attention and its correlation to crash potential	IQ questions and 1-back task.	75 participants (41 males), 21-69 with M=36.16 and SD=12.38.	Monetary compensation. Handout and two test drives.	University of Leeds Driver Simulator. FaceLab eye tracker	Gaze behaviour and driving simulator data.
(Zeeb et al., 2016)	Effect of visual and cognitive distraction of-of takeover quality	Naturalistic tasks: reading news, writing an email, and watching a video.	79 participants (44 males), 35-45. Participants were separated into 7 groups.	A demographic questionnaire, consent form and 4 test drives. Also, training on the multimedia device used in the experiment. Monetary	Mercedes Benz driving simulator. Two video cameras	time to eyes on, time-to-hands-on, time-to-system-deactivation, deviation from the centre of the lane, lateral acceleration after the takeover.
(Gold et al., 2013a)	Quantify adequate , time for a driver to perform a takeover.	Surrogate Reference Task	32 participants (24 males), 19-57, M=27.6, SD=8.7	Questionnaire, briefing and test drive.	High fidelity driving simulator. 3 cameras and Dikablis eye tracking	Gaze behaviour and simulator data.

(Wright et al., 2016a)	Assess takeover time for experienced drivers.	Naturalistic task: Reading task on on an iPad	36 participants. A young group with M=20.3 and middle-age group with M=37	Questionnaire, Briefing and test drive.	Real-time Techn-ologies Inc driving simulator. Mobile Eye	Gaze behaviour and simulator data.
(Merat et al., 2012)	Effect of secondary tasks on take over time and performance	TQT	50 participants, 28-68 M=47.38, SD=10.37	Questionnaire, Briefing on simulator, secondary task and test drive.	University of Leeds Driver Simulator. FaceLab eye tracker.	Gaze behaviour and simulator data.
(Larsson et al., 2014)	Effect of ACC familiarity on takeover quality	N/A	31 participants (24 males). Two groups - no ACC experience (M=38) and with ACC experience (M=55).	Questionnaire, Briefing and test drive.	VTI Driving Simulator III	Speed reduction, brake response time.

2.4 Detection of Driver Inattention

As identified by Singh (2015), driver inattention is responsible for over 48% of observed vehicle accidents in manual driving. Driver inattention is a general term describing both driver's fatigue and distraction in general (Craye and Karray, 2015; Regan et al., 2011). Driver distraction is defined and reviewed in section 2.3.2 and driver fatigue is defined and reviewed in section 2.3.3.

In the following sections, the literature of automatic detection methods of each type of driver inattention is critically reviewed. The following subsections discuss the different approaches adopted by the literature to detect driver inattention. As previously stated, methods for detecting fatigue, visual distraction, cognitive distraction, manual distraction are reported. Finally, a summary is presented at the end of this section.

2.4.1 Detecting Fatigue

Several approaches have been used to detect driver fatigue and cognitive distraction. Drivers are less accepting of invasive methods than non-invasive ones (Barr et al., 2009; Seppäläinen and Landrigan, 1988). Invasive approaches require the attachment of sensors or devices on the participant's body while non-invasive approaches use external sensors.

Ibrahim, (2014) split the methods of driver's fatigue detection into five main categories based on their measurement techniques. Firstly, physiological measurement uses physiological elements and quantifies their changes to estimate fatigue (e.g., Dinges and Grace, 1998; Thomas *et al.*, 2015). Secondly, physical activity measurement uses sleep and consciousness patterns to detect fatigue (e.g., Anderson *et al.*, 2017). This requires around the clock data collection to detect general fatigue patterns for the participants. This approach has more applications in aviation and is unsuitable for an automotive scenario because detection of driver's patterns will require an invasive data collection of driver's lives before and after driving sessions. Thirdly, behavioural measurement, which is the monitoring of driver's responses to tasks being carried out (e.g., Mabbott, 2003). These tasks could be either driving tasks or other tasks such as pressing a button when an action is required. Fourthly, mathematical models, which are models used to predict sleeping cycles and deprivation using data such as work times, sleeping hours, types of tasks performed and so on (e.g., Dongen, 2004). Finally, some hybrid techniques merging more than one of these measurements also exist in the literature (e.g., De Rosario *et al.*, 2010). In this section, physiological measures are discussed because it is the only category that fits with the limitations of the automated driving systems.

The physiological measures used in the literature were brain- heart- and eye-related measures. For example, ECG and EEG measures were used to assess fatigue. "An electrocardiogram (ECG) is a test which measures the electrical activity of your heart to show whether it is working normally" (NHS, 2015a). EEG is "a recording of brain activity. During the test, small sensors are attached to the scalp to pick up the electrical signals produced when brain cells send messages to each other.", (NHS, 2015b). Li *et al.*, (2011) identified that the most reliable and objective measure of driver fatigue is using electroencephalograph (EEG) and associated brain wave activities. Craye and Karray, (2015) has used both ECG and EEG signals to infer driver's fatigue level using a Support Vector Machine to classify four different states of fatigue. The study has shown accuracy between 87% and 93% for the different states. It is worth noting that while EEG and ECG are highly accurate, their invasiveness and the difficulty to deploy them in modern vehicles and driving simulator experiments make them an unpreferable choice in industrial research.

Eye activities are another significant factor in estimating fatigue. Studies since the 90s have shown their great potential (Stern et al., 1994). Eye activities include blinking behaviour, eye closure, pupil size, eye movement patterns and the percentage of eye closure (PERCLOS). Each technique provides a strong indication of fatigue signs. Most of the literature has shown that PERCLOS is the main feature for fatigue detection (Masala and Grosso, 2014). PERCLOS is a psychophysiological measure that indicates a person is fatigued when their eyes are closed for more than 80% of the time during a certain period (Dinges and Grace, 1998). Typically, it is the primary technique used in commercial fatigue detection devices (Ibrahim, 2014).

In addition, the eye blinking rate is another method to detect fatigue. Both the rate and duration of eye blinking were extensively studied (Caffier et al., 2003; Lal and Craig, 2002). In these two studies, both rate and duration had a proportional relationship with fatigue level. This shows an interesting factor to detect the level of fatigue. To measure the duration of the blink, Senaratne et al. (2011) used the optical flow technique achieving 82.7% of accuracy. The optical flow technique tracks the velocity of the upper eyelid through image comparison. Ibrahim, (2014) noted that the optical flow model should be combined with the eye closure measurement for more reliable detection since blink detection is sensitive to the lighting environment.

Another critical factor is pupil size. Interestingly, it had a strong relationship with the blink rate (Nakayama, 2006) and was found that the pupil width decreases as fatigue increases (Morad et al., 2000). Nishiyama et al. (2007) used a high-resolution infrared camera to accurately detect the width of the pupil in a driver fatigue detection scenario. While this gives another indication, IR cameras required the installation of extra hardware and showed some malfunction during daytime measurements (Hartley et al., 2000).

Table 3: Types of Visual Distraction Metrics (Klauer et al., 2006).

Eyes off forward roadway metric	**Definition**
Total time eyes off the forward roadway	The number of seconds that the driver's eyes were off the forward roadway during the 5 seconds prior and 1 second after the onset of the precipitating factor
Number of glances away from the forward roadway	The number of glances away from the forward roadway during the 5 seconds prior and 1 second after the onset of the precipitating factor
Location of longest glance away from the forward roadway	The location of the longest glance (as defined by the length of the longest glance) , location is based upon distance (in degrees) from centre-forward and is in one of three categories: <15°, between 15°and 30°, >30
Length of longest glance away from the forward roadway	The length of the longest glance that was initiated during the 5 seconds prior and 1 second after the onset of the precipitating factor

2.4.2 Detecting Visual Distraction

Visual distraction is straightforwardly detected using eyes-off-road glances. Klauer et al. (2006) used eyes-off-road metrics to evaluate crash or near-crash cases of one hundred naturalistic car driving and identified the metric described in Table 3. The findings of the study indicated that eyes-off-road model is time dependent as the "length of eye glance from the forward roadway increases, the odds of being in a crash or near-crash also increases ... Risk percentage calculations suggest that 23 per cent of the crashes and near-crashes that occur in a metropolitan environment are attributable to eyes-off-the- forward-roadway greater than 2 seconds", (Klauer et al., 2006).

Detection of the eyes-off-road feature is done using either a single camera vision system or eye-tracker devices. The eye-trackers could be either glasses worn during the experiment or a desktop bar with multiple cameras. A few examples of such devices are Tobii Glasses (Tobii, 2016a), Dynavox (Tobii, 2016b) and FaceLAB eye-tracker (faceLAB, 2016). These devices are more commonly used than the computer vision-based solution because of their higher accuracy and simplicity. Alternatively, computer vision-based solutions use RGB or RGB-D cameras to detect eyes movements using mathematical algorithms using single or multiple cameras. Interestingly, the eye trackers method is the most commonly used in the literature of human factors in highly automated driving studies (Körber et al., 2015a; Merat et al., 2014, 2012).

2.4.3 Detecting Cognitive Distraction

Cognitive distraction (or called mind-off-road, mental distraction) is a mental distraction that happens when the driver is deeply involved in thoughts other than driving tasks and safety. Some studies showed that physiological behaviour provides good accuracy in detecting cognitive distraction. For example, a study concluded that ECG is the most sensitive measure to mental workload (Paxion et al., 2014). Using n-back task, Reimer et al. (2009) imposed an incremental workload on participants with 0-back to 3-back tasks. Results have shown that heart rate increases step wisely (see Figure 7) as the cognitive workload increases. The study has also shown that skin conductance increased proportionally with the cognitive workload. Another non-invasive approach used Support Vector Machines (SVM) to predict cognitive distraction with an accuracy of up to 96% using eye tracking features (Liang et al., 2007). SVM is a discriminative classifier that uses labelled training data for classifying data categories (Suykens and Vandewalle, 1999).

Moreover, a novel approach was implemented using ECG to detect driver's distraction. Fifteen drivers between twenty and fifty years old were asked to drive with and without a secondary task on a single lane road on a driving simulator. The secondary task was double-digit addition arithmetic task. Results showed a significant difference between the multi-scale entropy of the ECG signal that could be linearly classified (Yu et al., 2011).

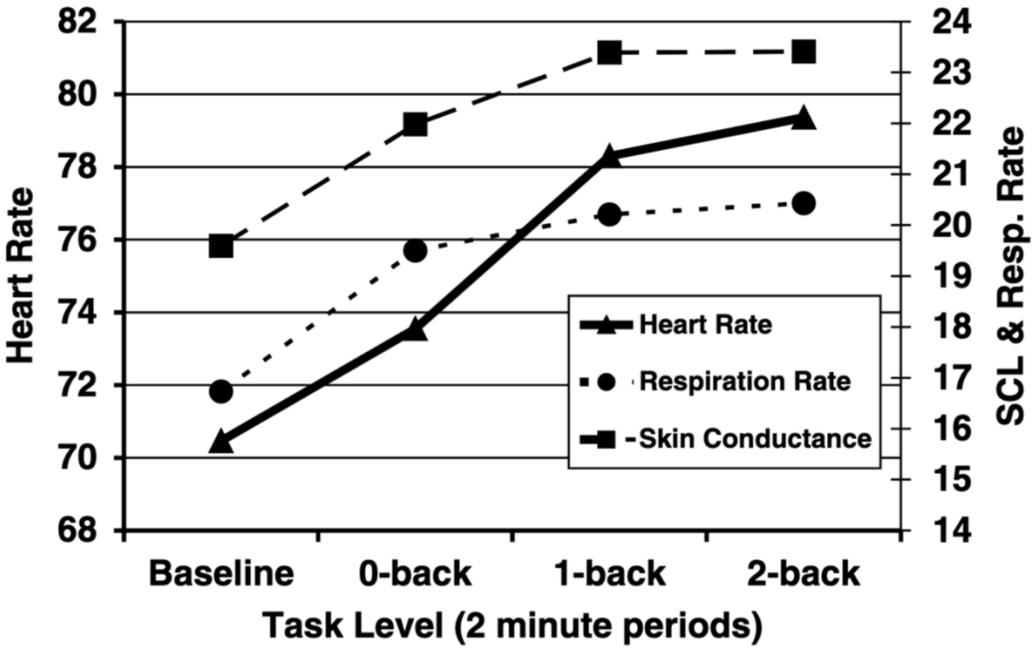


Figure 7: An illustration of how heart rate increases as the number of variables in the n-back task increases. This signifies an increase of the heart rate as the cognitive workload of participants increases. The n in the n-back task refers to the number of the variables that participants had to retain in their memory during the n-back task (Reimer et al., 2009).

ECG is sometimes used to aid and support the accuracy of studies. A study has used a stereo camera to track both head and eyes movements, and pupil size features in addition to ECG and heart rate to aid and test the accuracy of the algorithms. Support Vector Machine and Adaboost were used to fusion these features. Drivers were distracted using conversational and arithmetic tasks. The conversational task was to describe a road they usually commute on, and the arithmetic task was to subtract seven from 1000 consecutively. Results have shown an average of 86% accuracy in detecting cognitive distraction (Miyaji et al., 2009).

Table 4: A comparison of different machine learning algorithms used to classify cognitive distraction (Fernández et al., 2016).

Algorithm	**Feature**	**Classifier**	**Accuracy**
(Zhang et al., 2004)	Eye gaze-related features and driving performance	Decision Tree	81
(Zhang et al., 2004)	Eye gaze-related features	Decision Tree	80
(Zhang et al., 2004)	Pupil-diameter features	Decision Tree	61
(Zhang et al., 2004)	Driving performance	Decision Tree	60
(Liang et al., 2007)	Eye gaze-related features and driving performance	SVM	83.15
(Liang et al., 2007)	Eye gaze-related features	SVM	81.38
(Liang et al., 2007)	Driving performance	SVM	54.37
(Liang et al., 2007)	Eye gaze-related features and driving performance data	DBNs	80.1
(Miyaji et al., 2010)	Heart rate, Eye gaze-related features and pupil diameter	Adaboost	91.5
(Miyaji et al., 2010)	Eye gaze-related features	SVM	77.1 (arithmetic task)
(Miyaji et al., 2010)	Eye gaze-related features	SVM	84.2 (arithmetic task)
(Miyaji et al., 2010)	Eye gaze-related features	Adaboost	81.6 (arithmetic task)
(Miyaji et al., 2010)	Eye gaze-related features	Adaboost	86.1 (arithmetic task)
(Yang et al., 2015)	Eye gaze-related features and driving performance data	ELM	87
(Yang et al., 2015)	Eye gaze-related features and driving performance data	SVM	82.9

Wearable technologies at a consumer level achieved a reasonable accuracy in detecting the cognitive workload using collected physiological data of the driver (i.e., heart rate) based on Melnicuk's *et al.*, (2016) study. Their experiment took place on a simulator with 14 participants driving in rural, motorways and other roads and with a simulated accident. Results showed reliability in detecting cognitive workload based on heart rates (Melnicuk et al., 2016). However, Wang et al. (2016) compared a set of wearables against ECG and recommended that "Electrode-containing chest monitors should be used when accurate heart rate measurement is imperative". Polar H7, a chest-based wearable to monitor heart rate, achieved a 99% accuracy and Apple Watch was next with an accuracy of 91%. Even though Wang et al. (2016) participants were limited to a set of young healthy participants, it still gives a good indication of the accuracy of wearables. The findings also align with the findings of Melnicuk et al. (2016).

Also, a smartwatch was used in a full system for assessing driver's vigilance as a processing unit for non-invasive ECG and PPG fabrics attached to the driving wheel (Lee et al., 2016). PPG is a plethysmogram that is obtained optically to detect blood volume changes. Lee's et al., 2016 study found out that using both gender and age of the driver in the model increased the accuracy of the classifier. Using machine learning methods, the probability of driver's distraction level was calculated, and then a warning through the smartwatch is issued when the vigilance level fell below a threshold. The study showed an approximate 97.28% accuracy (Lee et al., 2016).

Fernández *et al.*, (2016) performed a review of different machine learning methods and the features used in different cognitive distraction studies. A summary of their review is illustrated in Table 4. Most of their reviewed studies used eye and gaze-related features and pupil diameters during the driving task to predict cognitive distraction. Several classifiers were used such as Decision Trees, Support Vector Machines, Adaboost and others, more information about machine learning methods is illustrated in section 2.6.

2.4.4 Detecting Manual Distraction

Manual distraction is concerned with the arm posture of drivers. Manual distraction has limited research in the manual driving field; though, (Craye and Karray, 2015) proved that arm posture is the highest cue for driver distraction in their study. A Kinect camera was used to extract four arm postures and used Hidden Markov Model and AdaBoost classifiers to fusion PERCLOS, head pose, orientation, and expressions together. The Hidden Markov Model is a statistical model that observes previous states to predict how likely the next state would be. AdaBoost is a machine learning meta-algorithm that merges several other machine algorithms to improve their overall performance. The Kinect camera made it easier to classify since it adds a depth layer into the RGB images making it easier to extract features from data. The highest accuracy in the study was 89% at estimating driver distraction. Another study by Park and Trivedi, (2005) used body poses including driver's static pose, dynamic gesture, body-part action, and driver-vehicle interaction to build an activity recognition framework for driver's activities. While arm positioning was essential for the classification in this study, the trend of research in this field focuses more on facial features and gaze estimation.

2.4.5 Summary of Driver Inattention Detection

As discussed in Chapter 1, conditional automated driving systems (Level-3) have to identify its limitations using system boundaries (SAE International, 2018). For example, driving in construction sites or under heavy weather conditions may be challenging to the automated systems due to sensory limitations. When a system boundary is detected, a take-over request is issued to the driver to take over the vehicle's control. The take-over request (TOR) will have to be prompted in a timely manner (7 seconds for instance, (Gold et al., 2013a)) allowing the driver to perform a safe transition before a potential collision is expected (NHTSA, 2013). This time is usually spent by drivers in regaining situational awareness and planning a maneuverer in addition to motor readiness which may occur sequentially or in parallel (Zeeb et al., 2015). Motor readiness is identified as the time it takes a driver to regain mechanical control (i.e., hands-on wheel and feet on pedals) (Zeeb et al., 2015) while response time is the time between a TOR and the driver applying a significant change on braking or steering wheel (Zeeb et al., 2015). Several studies showed that motor readiness is consistent, ranging from 1.2 to 1.8 seconds (Gold et al., 2013a; Zeeb et al., 2015).

Many studies examined the influence of secondary tasks on gaze related measures during highly automated and manual driving (Marquart et al., 2015). In previous studies, takeover time and performance have been correlated with eye blinking (Merat et al., 2012), gaze behaviour (Gold et al., 2016; Ko and Ji, 2018; Louw et al., 2016; Wright et al., 2016a; Zeeb et al., 2016, 2015), eye movements and PERCLOS (Jamson et al., 2013). Two manual driving studies induced mental workload using verbal and spatial imagery tasks (Recarte and Nunes, 2003, 2000). Their results indicated that their NDR tasks caused pupil dilation which indicated a high mental workload. Finally, the literature had limited to no studies that investigated pupil diameter changes in a highly automated driving environment. This literature limitation is observed, and our study decided to investigate the effect of distraction in HAD on the pupil diameter of drivers.

However, few studies examined the effect of NDR tasks on heart rate (Carsten et al., 2012; de Waard et al., 1999; Wille et al., 2008). The heart rate measured in highly automated driving is lower than manual driving and ACC driving

(Carsten et al., 2012) which matches with findings of de Waard *et al.*, (1999) that reported a slight decrease (73.2 vs 74.0 beats/min) in heart rate during automated driving. This difference is an indication of mental workload reduction (De Winter et al., 2014). Moreover, in manual driving studies, heart rate increased incrementally with increasing mental workload, and a plateau of the physiological measures was observed (Mehler et al., 2009). Moreover, the study reports a decrease in driving performance as the mental workload increases. Wearable technologies at a consumer level achieved a reasonable accuracy in detecting the cognitive workload using collected physiological data of the driver (i.e., heart rate). In the study by Mehler et al (2009) the experiment took place on a simulator with 14 participants driving in rural, motorways and other roads and with a simulated accident. Results showed reliability in detecting cognitive workload based on heart rates (Melnicuk et al., 2016). Thus, the study adopted measuring heart rate to identify cognitive distraction of drivers in a HAD environment.

2.5 Performance Measures of the Takeover

The driver's performance in highly automated driving studies was assessed using several methods (Radlmayr et al., 2019). Several studies used minimum time-to-collision (min-TTC) (Gold et al., 2016; Körber et al., 2016; Radlmayr et al., 2014), longitudinal acceleration (Gold et al., 2016; Radlmayr et al., 2014), braking (Körber et al., 2016; Larsson et al., 2014), minimum speed (Larsson et al., 2014) and occurred collisions (Radlmayr et al., 2014; Wandtner et al., 2018) to assess the quality of Drivers' Performance . Such scarcity makes it challenging to provide a cross-comparison among studies. Motivated by these limitations, Radlmayr et al., (2019) reported new take-over performance measures named Takeover Performance Measures (TOPS) that aggregates vehicle, mental and subjective ratings of the take-over to provide a single metric assessing the takeover. Though, the study has not provided any correlation between driver's physiological changes and the TOPS results for each participant.

Based on Radlmayr et al., (2019), Drivers' Performance measures could be split into the following categories:

1) Driver-related: reaction time, eyes-on-road time, etc.,
2) Vehicle-related: braking, time-to-collision, acceleration, etc.,
3) Subjective measures: usually collected through a questionnaire at the end of the experiment. A breakdown is illustrated in Figure 8 to show more examples of the aforementioned three types of performance measures.

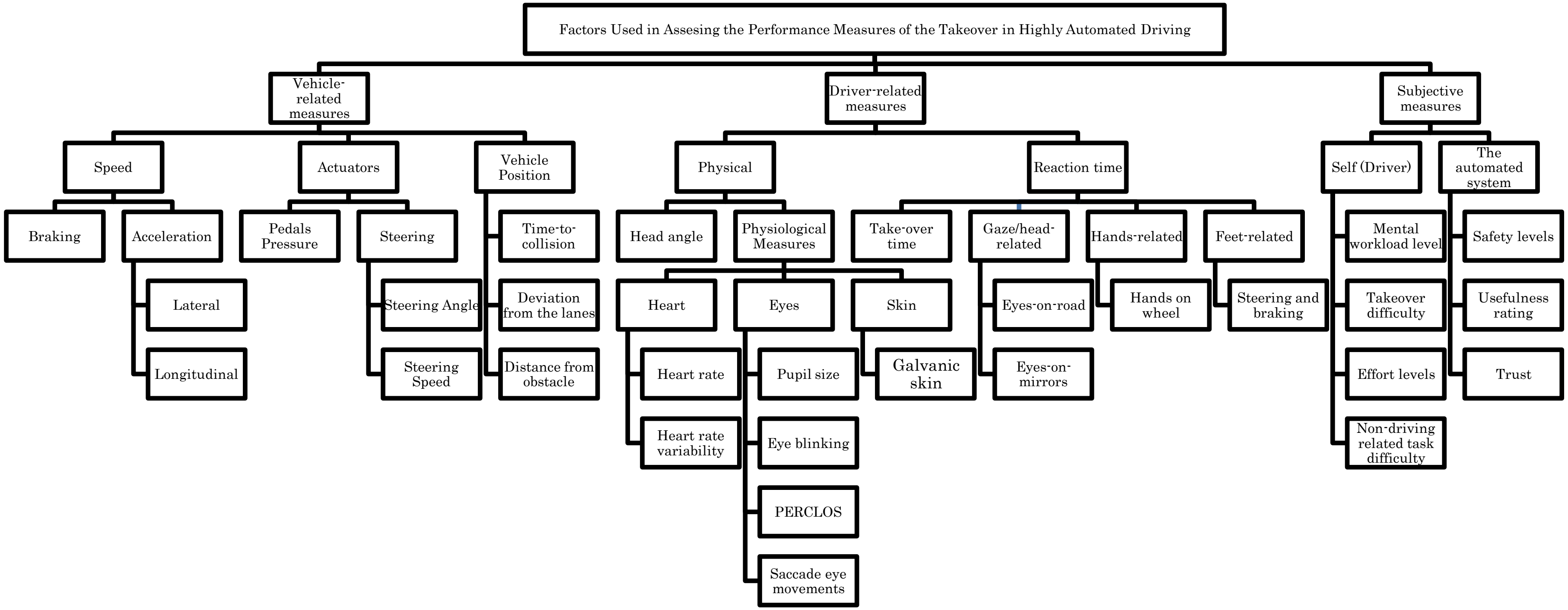


Figure 8: A breakdown of the performance measures used by several HAD studies. The wide range of those performance measures and their interdependency make it difficult to compare results among studies. The data is gathered through the literature review of Radlmayr et al., (2019).

Driver-related measures are variables measuring a reaction or a physical move made by the driver. For example, take-over time is one of the most prominent objective measures for evaluating the take-over performance (Eriksson and Stanton, 2017; Gold et al., 2016, 2013a, Zeeb et al., 2016, 2015). In addition, eyes-on-road time is another objective measure which is defined as the time it takes a driver to fix their eyes on the road, probably after being distracted by a visual NDR task (Gold et al., 2013a; Vogelpohl et al., 2018) in addition to steering and braking reaction times (Eriksson et al., 2017; Happee et al., 2017). To conclude, driver-related performance measures are time-based and are an indirect measure of a driver's mental state. Those performance measures are easy to compare to but difficult to accurately collect because experimenters have to perform video labelling to measure the start and end of each event (Zeeb et al., 2015).

The driver-related measure covers their physiological behaviour. Gable *et al.*, (2015) used heart rate and pupil size as objective physiological measures of mental workload. Their results showed that heart rate and pupil size were valid objective measures of the workload in manual driving and may require a small set of participants to produce a valid dataset. However, physiological measures may have an indirect influence caused by the non-driving mental or physical stimuli (Teh et al., 2014). It's argued that physiological data must be normalised to cancel individual differences among the collected data; thus, allowing better comparison among them (Cain, 2007).

Vehicle-based performance measures are variables measuring the movement or the change of vehicle's position, lateral or longitudinal speed. For example, the minimum time-to-collision (TTC) is a prominent vehicle-related performance measure (Radlmayr et al., 2019). Lateral and longitudinal acceleration, steering wheel angle and braking level are also popular measures. Most of the vehicle-related measures are easy to collect and compare, in comparison to physiological measures. Though, most of the HAD studies have not employed an effort in providing their rationale for their choice of vehicle-related measures or provided a detailed approach in replicating the same measure. The lack of such information makes it difficult to compare studies and find common ground for the choice of methodologies in HAD studies.

Finally, subjective measures are values collected through questionnaire or surveys during or at the end of the experiment to understand the driver's perceived understanding of the driving task. For example, subjective measures could be the driver's perceived workload during the takeover process (Eriksson et al., 2017), criticality rating (Naujoks et al., 2016) and finally, a rating of difficulty (Zeeb et al., 2016).

Subjective measures of drivers usually assess their mental workload indirectly by asking drivers to estimate the difficulty of the task; especially under repeated exposure to the same task (Teh et al., 2014). Many studies (De Winter et al., 2014; Ko and Ji, 2018) based their questions on the NASA-TLX method (Hart, 2012). A previous study by Gopher and Braune, (1984) showed that subjective measures have an accuracy of 0.9 or higher in unidimensional ratings. However, few HAD studies reported a disassociation between objective and subjective measures (Zeeb et al., 2016) which was reported in manual driving studies too (Horrey et al., 2009a). This disassociation shows a research gap that the study aimed at exploring.

2.6 Predicting Drivers' Performance Using Machine Learning

2.6.1 Introduction

"Machine learning is programming computers to optimize a performance criterion using example data or past experience.", (Alpaydin, 2014). The definition explains that machine learning produces computer programs that are able to solve a problem based on previous examples. A machine learning algorithm uses input and output data to train on. Once the training is done, the ML algorithm is able to find a solution to any input data providing that the learning phase was successful (Alpaydin, 2014).

Machine learning (ML) has been extensively used in manual driving studies to assess situational awareness and fatigue (Sikander and Anwar, 2018). For example, several ML methods were used to assess situational awareness such as Bayesian probabilities (Armand, 2016) and Support Vector Machine (Solovey et al., 2014). To identify the driver's inattention, studies used a wide variety of input data to train the ML models. Darzi *et al.*, (2018) identified two categories of input data used in machine learning studies concerning driver inattention; vehicle-based and physiological based data.

2.6.2 Input Data

The physiological data were the dominant data used to build up many machine learning models. For example, physiological data were used as an input to assess mental workload (Solovey *et al.*, 2014b; Sikander and Anwar, 2018), vigilance (Hecht et al., 2019) and fatigue (Sikander and Anwar, 2018). The first known study that assessed physiological behaviour of drivers to automatically classify their hazarded state were (Healey and Picard, 2005) using a range of sensors including galvanic skin sensor, ECG and EEG to classify stress levels of drivers in manual driving scenarios. Since then, many studies investigated physiological behaviour of drivers to assess their mental workload (Brookhuis and De Waard, 2010), stress level (Wijsman et al., 2011) and distraction (Hirayama, et al 2016).

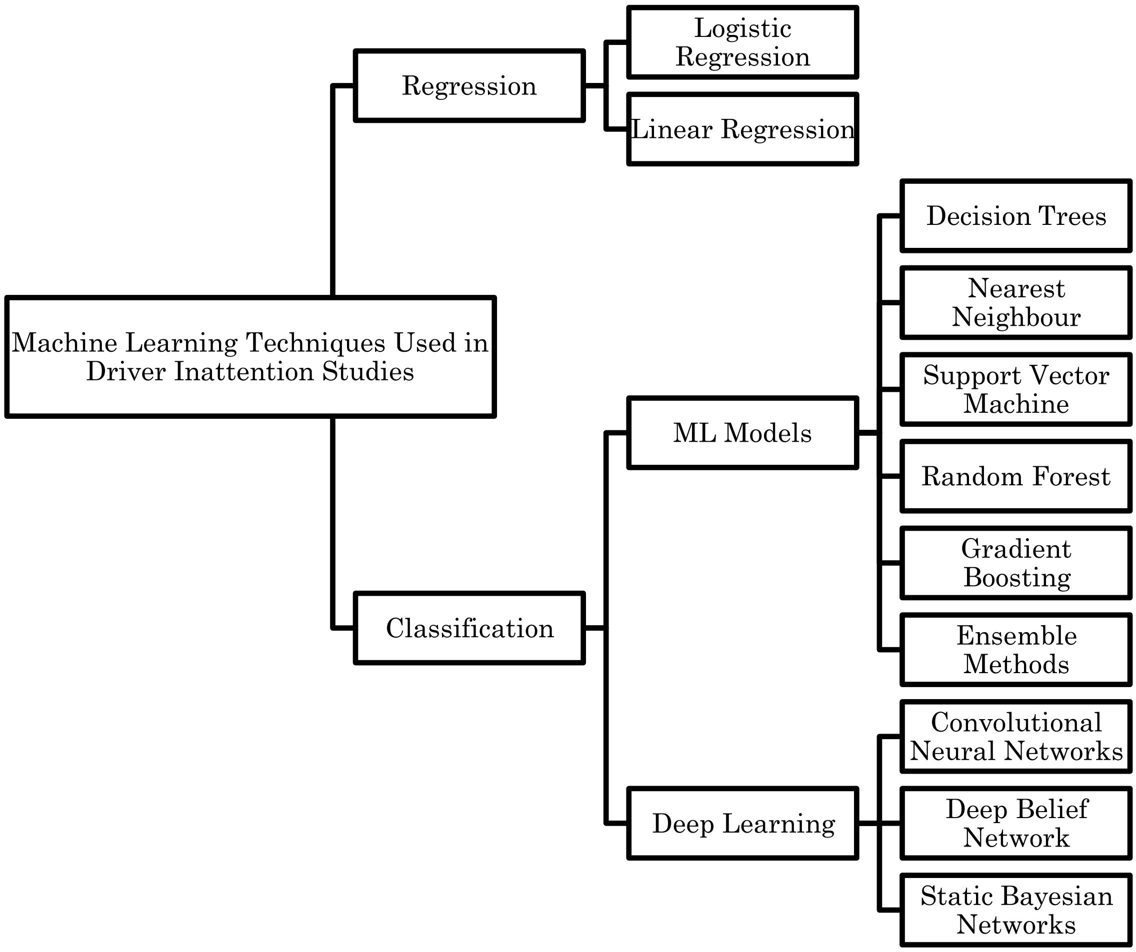


Figure 9: A breakdown of all machine learning techniques used in driver inattention studies. The techniques are broadly split into two categories: regression and classification. The classification approach is the most widely used approach with a wide variety of classifiers. Recent studies adopted different types of deep learning neural networks to identify driver inattention; however, this approach is not extensively surveyed since it's out of the scope of the study.

The vehicle-based approach uses vehicle metrics to identify distracted drivers using an indirect approach of analysing their driving moves such as lane keeping or lane changes (Harvey et al., 2011). This approach uses vehicle-based signals such as vehicle kinematics, steering angle change and other factors to assess the driver's interaction with the vehicle (Choudhary and Velaga, 2017; Zheng and Hansen, 2016). This approach achieved an accuracy of 98.7% in a study that used 8 signals signifying the interaction of the driver such as pedals pressure, steering wheel angle and engine RPM (Jafarnejad et al., 2018). A major drawback to this approach is its inapplicability to the highly automated driving environment where drivers don't interact with the vehicle during the automation phase.

So far, the input used in the machine learning methods have been discussed and in the next few paragraphs different machine learning techniques used in the driver inattention detection studies will be discussed.

2.6.3 Machine Learning Models

Machine learning techniques used in driver inattention studies could be broadly categorised into shallow and deep learning models. Shallow models represent a group of classifiers that are easy to train, understand and provide reasonable accuracy (Sikander and Anwar, 2018). For example, Support Vector Machines are extensively used in several driving inattention detections studies (Liang et al., 2007; Miyaji et al., 2009; Solovey et al., 2014) in addition to logistic regression (Darzi et al., 2018), ensemble methods (Miyaji et al., 2009; Zhang and Hua, 2015), gradient boosting (Hu and Min, 2018), Nearest Neighbour, Random Forest (Jafarnejad et al., 2018) and decision trees (Le et al., 2018). A breakdown of those methods is illustrated in Figure 9. A major drawback to the shallow models is their need to be fed a distinctive set of features. Consecutively, this requires a knowledge-based identification of the right set of features that domain experts could identify because the chosen features could potentially have an accurate prediction of the output (Wu et al., 2018).

The deep learning approach can extract and identify unique features inside the data to enable better classification. However, they require a large set of data and extensive training and validation (Sikander and Anwar, 2018). For example, Convolutional Neural Networks were the first models adopted in the driving inattention detection studies (Sikander and Anwar, 2018). While the deep learning approach is promising, it requires a large dataset to train models which is challenging to collect.

The literature has a minimal number of studies that covered the performance of the handover process. One study used psychometric tests to predict driver's takeover time in highly automated driving (Körber et al., 2015b) – the study used multitasking and reaction time tests before the driving to find whether such skills can predict the takeover time of drivers. The approach shown in Körber *et al.*, (2015) was inheritably limited since it has no feedback on driver's mental state at the takeover; making it limited and unreliable for real-world driving environment.

The literature review identified a gap in the use of machine learning techniques in highly automated driving to predict driver's performance. More work is required to identify whether the findings of manual driving studies could be applicable to highly automated driving field – specifically, whether machine learning techniques using physiological data as an input can predict driver's performance in highly automated driving.

2.7 Research Gap

In the past four years, several studies have identified the main vital issues that impaired the driver's performance during the handover process in Level-3, automated vehicles. Researchers have explored several ways to mitigate the deficit of performance through "stepped handover (Gold et al., 2013b), feedback systems (Lorenz et al., 2014) and trust in the reliability of automated systems to perform efficiently at all times (Beller et al., 2013b)" (Morgan et al., 2016: 12).

However, a system that predicts driver's future takeover time and performance using physiological measures collected right before a takeover request have not been investigated yet even though it was recommended (Wright et al., 2016a) as future work. In previous sections, takeover time and performance have been correlated with eye blinking (Merat et al., 2012), gaze behaviour (Gold et al., 2016, 2013a; Louw et al., 2016; Wright et al., 2016a; Zeeb et al., 2016, 2015), eye movements and PERCLOS (Jamson et al., 2013). Therefore, the hypothesis of this study assumes the possibility of predicting future takeover time and performance based on the listed features. Moreover, the Driver Controllability Set introduced by Nilsson, Falcone and Vinter, (2015), that was able to label a takeover process with safe or unsafe could provide a foundation for a feedback system for self-learning during the transition from Level-3 to Level-1. The proposed system would give the automated driving system an eye inside the car to better understand the driver's limitations and capabilities at all times to plan for appropriate takeover time and suggest an adequate feedback system suitable to driver's current state. Therefore, improving the safety level of automation will be potentially introduced.

The main aim of the study was to examine physiological changes caused by cognitive and visuo-cognitive secondary tasks and how they influence response time and quality during take-over scenarios. Reading and responding to an email and twenty questions tasks were reported to degrade response time and quality (Merat et al., 2012; Zeeb et al., 2016). Thus, the study investigated the influence of secondary tasks on physiological behaviour. Furthermore, the study was designed to examine whether the reported learning curve (Körber and Bengler, 2014; Larsson et al., 2014; Wright et al., 2016a) of take-over handling could have any effect on drivers' physiological changes before and during the takeover and how they correlate with response time and quality after a TOR. Specifically, the literature used objective and subjective measures to quantify the quality of the takeover. Few studies identified a convergence between reported subjective and objective measures (Zeeb et al., 2016). Thus, this study was designed to assess whether physiological changes at TOR could provide more information to explain the aforementioned convergence. A further aim was to propose new performance measures to assess the quality of drivers handling vehicles.

Key Gaps:

- How driver's physiological data correlates to the driver's attention. Current studies showed that physiological data are affected by NDR tasks, but no studies investigated heart rate and pupil size.
- No studies have investigated thoroughly how physiological data are affected by the driver's involvement with NDR tasks and the state of the automated driving system. Few studies compared driver's physiological data during their involvement in NDR tasks before the takeover, during the takeover and after handling a critical hazard.
- The literature shows an unaligned and unstructured approach in choosing takeover performance measures. The literature requires structured and well-implemented performance measures to standardize the results. This, in turn, will enable better comparison among studies.
- No studies were found to use physiological data as an input to predict driver's performance using machine learning techniques in a highly automated driving field during the take-over process. More work is required to assess and choose the right set of features to be used for the machine learning models.
- Further work is required to assess the suitability of the machine learning approach to predict driver's performance during the takeover in highly automated driving.

2.8 Summary

The literature review chapter discussed highly automated driving studies relating to driver's performance in the transition from automated to manual driving. This literature review concluded that there was a limited understanding of the driver's physiological behaviour in highly automated driving. Several human factors studies observed the human-machine interaction between the driver and the automated system. Researchers suggested that the human driver is turning more into an operator rather than a driver; however, there was no clear understanding of how drivers may recover from a distraction to take back the driving task to handle a potentially critical hazard during the transition phase.

Most HAD studies based their hypotheses on manual driving and flight control literature to discover the new aspects of how the human driver (operator) will handle the automated system. The literature review showed that the driver's inattention caused by NDR tasks would have the highest impact on the driver's performance during the transition period. Many studies identified a lack of situational awareness due to driver's involvement in tasks that demanded high mental and mechanical workload. While visual and mechanical workloads are easy to detect, the mental workload is the most difficult to detect and has potentially the highest impact on driver's performance; especially when combined with other visual or mechanical distractions. HAD studies showed conflicting results on whether NDR tasks with a demanding mental workload could enhance the driver's performance during the transition. Precisely, the impact of mental workload on drivers in HAD studies is a new phenomenon that hasn't been explored, and there is no equivalent phenomenon in either manual driving or flight control studies. Therefore, more work is required to understand the physiological cues imposed by mentally demanding NDR tasks and how they may correlate to the driver's behaviour in the vehicle. Moreover, there is limited research in understanding whether physiological measures could have a direct correlation with the driver's performance during the transition. Studying this correlation may open the door to allow the automated system to evaluate a human's vigilance and the ability to handle a critical hazard in any emergency takeover.

The research gaps listed in section 2.7 illustrate that physiological behaviour might give strong predictability power of the driver's performance in a critical hazard during the transition. Limited research was found in highly automated driving that explored driver's physiological data, specifically heart rate, eye movements and pupil size. The study's scope is to focus on physiological data that could be collected in production vehicles non-intrusively. However, non-intrusive approaches are usually less accurate, and their inaccuracies may bias the data analysis of the study. Therefore, the study aimed at using intrusive devices to collect those physiological data to ensure minimal error rate and to maximise the accuracy of the analysis. Those models may be deployed to the vehicle in the future, and more work will be required to assess non-intrusive approaches to collect the same data, but that is out of the scope of the study.

In addition, the research gap identified that assessing driver's performance had a non-standardised approach in the field of HAD. Therefore, the study seeks to introduce performance measures that 1) are predefined, 2) could be reused in other studies and finally 3) correlate with the driver's physiological behaviour. Even though many studies use time-to-collision and reaction time as the main metrics to identify driver's performance, the literature showed that driver's reaction time is not correlated to driver's ability to handle a critical hazard during the transition phase. Drivers were observed to react faster at the cost of skipping some safety-critical steps such as mirror checks whilst performing their collision avoidance manoeuvres. Therefore, more efforts are required to assess the driver's performance based on their ability to handle the vehicle. Some studies introduced relevant performance measures such as braking and lane changes but failed to introduce the mathematical formulas to reproduce those measures. Consequently, the literature needs new performance measures that are easy to reproduce, meaningful and correlate with the driver's physiological behaviour.

Finally, the prediction of driver's performance based on their physiological data hasn't been explored in HAD studies. Though, the same method showed great success in manual and flight control studies. Several studies used machine learning methods to classify driver's vigilance, workload and performance automatically. Hence, assessing the predictability of driver's performance based on their physiological behaviour using machine learning could enhance the communication loop in the human-machine interaction in HAD environment.

3 CHAPTER

Research Methodology

The literature review has shown that there's a possible correlation between drivers' responses and their physiological measures in highly automated driving environment. This study assesses this hypothesis as stated in Chapter 2. The following chapter covers the research methods used to conduct the research and explain the rationale behind the chosen approaches. The study delimitation is addressed at the end of the Chapter and a justification for each choice is presented in detail.

3.1 Research Philosophy

Oates (2006) stated that Research Philosophy is "the creation of new knowledge using an appropriate process, to the satisfaction of the users of the research". Therefore, selecting an appropriate process is needed to evaluate the outcomes of the study to ensure there is a quantitative correlation. To ensure the success of the project, proper communication of the 'big idea' is needed (Nightingale, 1984). Therefore, a Research Philosophy is considered the underlying assumption that defines the strategy of data collection, analysis and usage to satisfy the main aim while communicating an easy to understand description of work (Walliman, 2005).

There are two main major disciplines of research philosophies: positivism and interpretivism (Bryman, 2003). Table 5 compares the two major types of philosophies and their role in understanding reality. More details are laid out in the following subsections.

Table 5: Procedural Bases of understanding Reality (Cohen et al., 2013)

	Positivist	**Interpretivist**
Philosophical basis	Realism: the world exists and is knowable as it really is Organisations are real entities with a life of their own	Idealism: the world exists But different people construe it in very different ways Organisations are invented social reality
Theory	A rationale edifice built by scientists to explain human behaviour	Sets of meanings which people use to make sense of their world and behaviour within it
Methodology	Abstraction of reality, especially through mathematical models and quantitative analysis.	The representation of reality for purposes of comparison. Analysis of language and meaning
Basic units of reality	The collectivity: society or organisations	Individuals acting singly or together

3.1.1 Positivism

Positivism is the branch of philosophy that emphasises the observable and factual over the theoretical and metaphysical. Repetition of measurable phenomena in the experimental environment is a necessity in the positivism philosophy. In addition, the positivist hypothesis searches for factual knowledge of a phenomenon for the moment of time. Since time could change some facts, some researchers question the validity of the positivist hypothesis to ensure the study's objectivity and reliability over time (Bryman, 2003). Positivists believe the truth is observable through monitoring phenomena in an undisputed objective real world. An objective real world is an abstract world that doesn't incorporate any social values in neither understanding nor interpretation of its phenomenon since observations have to remain repeatable (Cornford and Smithson, 1996). Projecting such philosophy on automated driving studies concerning driver's behaviour, positivism is an acceptable approach for understanding common human interaction with the automation system. While social values may have an influence on drivers (Fleiter et al., 2010), a researcher should have a research approach that sets social values away from influencing driving behaviour.

3.1.2 Interpretivism

Walsham, (1995) states that "interpretive studies are aimed at producing an understanding of the context of the information system, and the process whereby the information system influences and is influenced by the context". Interpretivists argue that social values and facts cannot be separated. In other words, facts are interpreted on the interpretations of social values which disallows any replication of theories over periods of time as they are socially constructed and promoted (Cornford and Smithson, 1996). It's worth noting that this philosophy is associated with opinions, beliefs, feelings, and assumptions rather than scientific facts as Crotty, (1998) argued. Therefore, interpretivism is not useful with natural sciences as researchers in such fields observe phenomena that are disassociated with social understanding. Even though the proposed study takes the human factor in consideration, the outcomes of the study are expected to be scientific facts rather than socially constructed reality. Thus, the Interpretivism viewpoint is unsuitable for the study's aims and objectives.

Table 6: Research Approaches based on Cornford and Smithson (1996) list.

Philosophy	Approach	Method	Details
Positivist	Constructive	Frameworks	A prescribed sequence of events that can be applied when undertaking a particular piece of research e.g. evaluation of systems.
		Theorem proof	Use of formulae and procedures that can be applied to represent a problem.
		Prototyping	A typical instance of a solution is produced and can be tested before producing the final working version.
	Nomothetic	Laboratory experi ments	The researcher manipulates some variables and observes the results. However, the laboratory setting and identified relationships may not be applicable to real-world contexts.
		Field experiments	Experiments conducted in real organisations, thereby increasing the reality of results. However, few organisations are willing to undergo experiments, and it can be difficult to control variables.
		Surveys	Obtain views and practices at a single point of time and draw conclusions from the sample to the whole population using quantitative analysis techniques.
		Case studies	Phenomena are studied it its real-life context without interfering with the phenomena.
		Forecasting	Uses quantitative techniques such as regression analysis to provide insights into future events where variables may change, such as predicting the level of sales.
Interpretivist	Idiographic	Action research	Where the researcher participates with the subjects in the problem situation, rather than taking the role of the observer. The problem usually results in change. Data is collected during the participation to provide practical value to an organisation, and the researcher uses the data relate to theoretical knowledge.
		Case studies	An in-depth exploration of one situation, for example, to address the implementation of a new accounting system in a particular organisation. Yin defines the purpose of the method is to cope with the technically distinctive situation in which there will be many more variables of interest than data points.
		Ethnography	Where the researcher is immersed in an organisation and interprets the viewpoints of members of that setting
		Futures research	See forecasting.

3.1.3 The rationale for Choosing the Positivist Philosophy

The study's main focus was to observe the human's physiological changes influenced by their surrounding environment. Positivism is the natural sciences' central philosophy because it believes in the progression of knowledge and discovery. The positivists believe that there is a reality and science's aim is to discover it (Denzin and Lincoln, 2005). The objectives of the study aimed at exploring the common patterns in the driver's behaviour which is philosophically a 'reality' that science attempts to unveil. Thus, the positivist approach is used.

3.2 Research Approach

Research Approach is the structured approach used to collect data within a pragmatic perspective. Based on research objectives, a specific research approach is chosen and optimised for the type of data to be obtained, time and resources allocated for the research project (Cornford and Smithson, 1996). The research approaches could be constructive, nomothetic, or ideographic. See Table 6 for a brief analysis of the relationships among research philosophy, approaches, and methods.

The constructive approach uses models based on refined concepts, frameworks, and technical development to represent theories based on situations that couldn't be physically manifested. Even though such models could be observed theoretically through existing studies available on the literature, constructing such models could be easier through empirical observation (Cornford and Smithson, 1996).

The nomothetic approach uses quantifiable measures to collect and observe data empirically in order to realise a statistical judgement or prove a hypothesis (Cornford and Smithson, 1996). An example of the nomothetic approach is a formal mathematical analysis and lab experiments along with surveys. The primary goal of the nomothetic approach is to provide a generalised insight of collected data. Therefore, data collection and sampling should be done carefully. It's worth noting that this approach aligns well with the positivist philosophy and is suitable for this study presented in the thesis.

Quantitative data analysis is associated with the nomothetic approach. It's originally developed for natural sciences (Myers et al., 1997) to collect quantifiable data that could be measured and replicated precisely without refutation which allows hypotheses to be tested statistically (Cornford and Smithson, 1996). Quantitative analysis leads to a generalised understanding of results when collected data is large enough to accommodate for such an approach.

Finally, idiographic research aims at exploring and understanding a particular phenomenon such as cases or events within a specific context. Crotty (1998) argues that it is associated with human affairs and individual cases as identified as 'idiots'. Case studies or action research are the main methods used in this approach (Cornford and Smithson, 1996). Since the proposed study aims at proving a hypothesis irrelevant to its context, this approach is rejected.

Qualitative analysis is associated with idiographic research because it drives a deeper understanding of the issue by collecting more personalised and more opulent description and more profound level of information. Therefore, the outcome is more insightful than quantitative approaches. It's also associated with the interpretive philosophy as it aims at understanding reality within its surrounding values (Firestone, 1987).

To achieve the main aim of the study, data should be collected, analysed, and validated. Using the right philosophy, approach and methods enable for accurate and scientifically valid results. The study employs the positivist philosophy based on the nomothetic approach to achieve its aims. The quantitative approach is used to collect and analyse the collected data. A summary of approaches used in this study is represented in 3.3 with a brief justification for each choice.

The research specific approaches chosen for the proposed study are:

1. **Lab experiment**: this approach is used to simulate a driving scenario to collect quantifiable measures of the driver's behaviour before and during a takeover situation. In addition, it allows for comparing driver's behaviour during Level-1 and Level-3 automation. And finally, it enables repeatability of the study while maintaining safety for all participants.
2. **Questionnaire**: this approach is used to understand the driver's background such as age, gender, driving experience, etc., in addition to their confidence before and after the experiment on highly automated driving vehicles. The confidence of driver's over the automation level has been questioned by Larsson et al. (2014) and therefore, asking drivers about their understanding of the system before and after the system is required.
3. **System Development**: the study aimed at assessing the predictability of the driver's performance using machine learning models. Thus, the study used the Spiral development methodology to explore several machine learning models and explore different data preparation techniques, see sections 3.4.6 and 3.4.8 for more details.

3.2.1 The justification for the Use of the Lab Experiment Methodology

Real world driving comes naturally with risks, especially when using a premature technology such as highly automated driving. Lab experiments based on driving simulators have become an established approach in testing driver's experiences and abilities to handle the driving task that enables constructing a visual driving environment that's quite similar to normal driving (Kaptein et al., 1996). The main advantages of driving simulators are the flexibility to generate and replicate driving scenarios in a safe environment (De Winter and Happee, 2012) in addition to the safety of drivers when the driving scenarios involve risky manoeuvres or driving under stress (Brookhuis and De Waard, 2010) or alcohol (Banks et al., 2004).

There were some concerns among researchers on the impact of driving simulators on driver's behaviour because they provide a safe driving environment in comparison to real driving (Santos et al., 2005) – their study concluded that "a relative validity is attainable in most cases and absolute validity in some". The reported results align with Kaptein, Theeuwes and Van Der Horst, (1996). No studies were conducted to assess the influence of the driving simulator's environment on driver's behaviour in highly automated driving scenarios. However, most studies in the literature of highly automated driving have used a driving simulator, (e.g. Gold et al., 2013a; Körber et al., 2016; Zeeb et al., 2016, 2015).

Also, there were studies examining naturalistic driving in Tesla S models (Endsley, 2017). The aforementioned study run by Endsley, (2017) used an interpretivist approach to understand the issues of automation on the driver's mindset. She reported that her study lacked the controllability of the driving simulator's environment and argued that it might have affected her reported results. In addition, using a real driving environment comes with ethical concerns on the safety of participants and the replicability of the experiment.

Accordingly, this study decided to choose the driving simulator to create a reproducible driving scenario to limit the variables influencing the psychophysiological changes of drivers in order to find answers to the research questions raised in Chapter 1.

3.2.2 The justification for the Use of the Questionnaire Methodology

According to Zeeb, Buchner and Schrauf, (2016), there was an observed deviation between subjective and objective measures of driver's performance. In order to collect the driver's personal judgement of their takeover performance, a questionnaire was required. The questionnaire method was used in similar studies to collect driver's demographic information and their evaluation of their driving performance (Gold et al., 2016; Mok et al., 2017; Zeeb et al., 2016).

The study followed the same approach to collect driver's age, driving experience, their reported fatigue level during the experiment, automated driving experience, the difficulty rating of the three takeovers, the difficulty of NDR tasks. The questions followed the structure advised by NASA-TLX standards (Hart, 2012), see Appendix B for more details.

To conclude, the main motivation behind using the questionnaire was to collect the driver's personal performance measures of their interaction with the experiment. This is identified as the subjective measures of their performance – that's in order to study the correlation between the subjective and objective measures of their driving performance. In addition, the questionnaire provided a better understanding of drivers' demographics and ensured the recruited participants fitted within the study's inclusion and exclusion criteria. More information about participants inclusion and exclusion criteria is provided in section 3.3.2.

3.2.3 The justification for the Use of the System Development Methodology

In order to assess the predictability of the driver's performance, the study adopted the system development approach to build a machine learning model that yields the highest predictability rate. The development process of the predictive models is dependent on experimentations (Bramer, 2016). Thus, the study explored various software development methodologies (Summers, 2011). The study imposed a level of risk since the study's hypothesis was to assess the predictability of the driver's performance at the transition. The uncertainty of finding the right features nominated the Spiral Model because it's generally preferred when failure is probable (Summers, 2011). The study chose the Spiral model because the Spiral model's main focus is on reducing risk during the development process. This is done by splitting the project into small segments and to provide continuous evaluation of the risk levels of each decision.

3.3 Experiment Design

3.3.1 Participants

Data represented in this study were collected in an experiment run at Loughborough University Design School. There was a total of thirty-six participants recruited for this study (53% females) and between 20 and 30 years of age (M=25.8, SD=5.7). Participants were invited to the lab for a 90 minutes session that involved both training and experimentation. Participants were asked to fill in a demographics survey and a questionnaire regarding their driving experience.

Drivers had a minimum of two years of driving experience. In addition, they were required to have a normal or corrected-to-normal vision. Of all participants, 84.6% had no experience with advanced cruise control. Informed consent was obtained after the experimenter explained the required tasks approved by the university ethics committee panel, see section 3.3.7 for more details. All experimental procedures were conducted in accordance with the ethical guidelines of the hosting university.

Out of the 36 participants, three participants were excluded from the study for the following reasons. Two of these were due to some missing data during the data collection process and another participant who reported they were profoundly fatigued during the experiment in the questionnaire at the end of the experiment.

3.3.2 The justification for the Chosen Participants' Demographics

The study recruited participants between 20 and 30 years of age with a minimum of two years driving. The study focused on this age group because the literature showed that they had the highest adoption rate of automated driving systems among different age groups (Payre et al., 2014). Since previous studies indicated that driving experience was a factor in the driver's performance of handling automated driving systems, the study ensured that drivers had to have a minimum of two years driving experience (Larsson et al., 2014).

The number of participants was chosen based on similar studies in the human factors field (e.g., Gold et al., 2013a; Wright et al., 2016a), who recruited 32 and 36 participants and had the same number of independent variables. In addition, research studies that focused on physiological measures of drivers recruited 19 and 21 participants (e.g., (Darzi et al., 2018)). Thus, the study decided to recruit a total number of thirty-six participants.

Participants who reported a 3-5 fatigue level on the NASA-TLX questionnaire at the end of the experiment were excluded from the data analysis, see Appendix B for more details. The subjective fatigue level of drivers was taken into consideration due to several studies reporting a significant impact on the psychophysiological changes of the drivers caused by their fatigue level (Lal and Craig, 2001). While fatigue is one of the variables affecting the driver's performance as indicated in section 2.3.3, the study has not considered it to avoid the potential of high dimensionality in addition to their detrimental impact on the clarity of the results. The study discusses the delimitations of the experimental design in section 3.5.

In addition, to ensure participants' alertness level was not affected, the study was only conducted twice a day. The first session was between 11 and 12:30 AM and the second one was from 4:00 to 5:30 PM. Picking these two timeslots was motivated by the results of Kraemer *et al.*, (2000) that showed that humans had two peak alertness level at 11:00 AM and at 3:00 PM. The first timeslot was centred at the first peak; however, the second one was one hour further than the second peak. The author decided to choose 4:00 PM for the second slot because it was more convenient for participants.

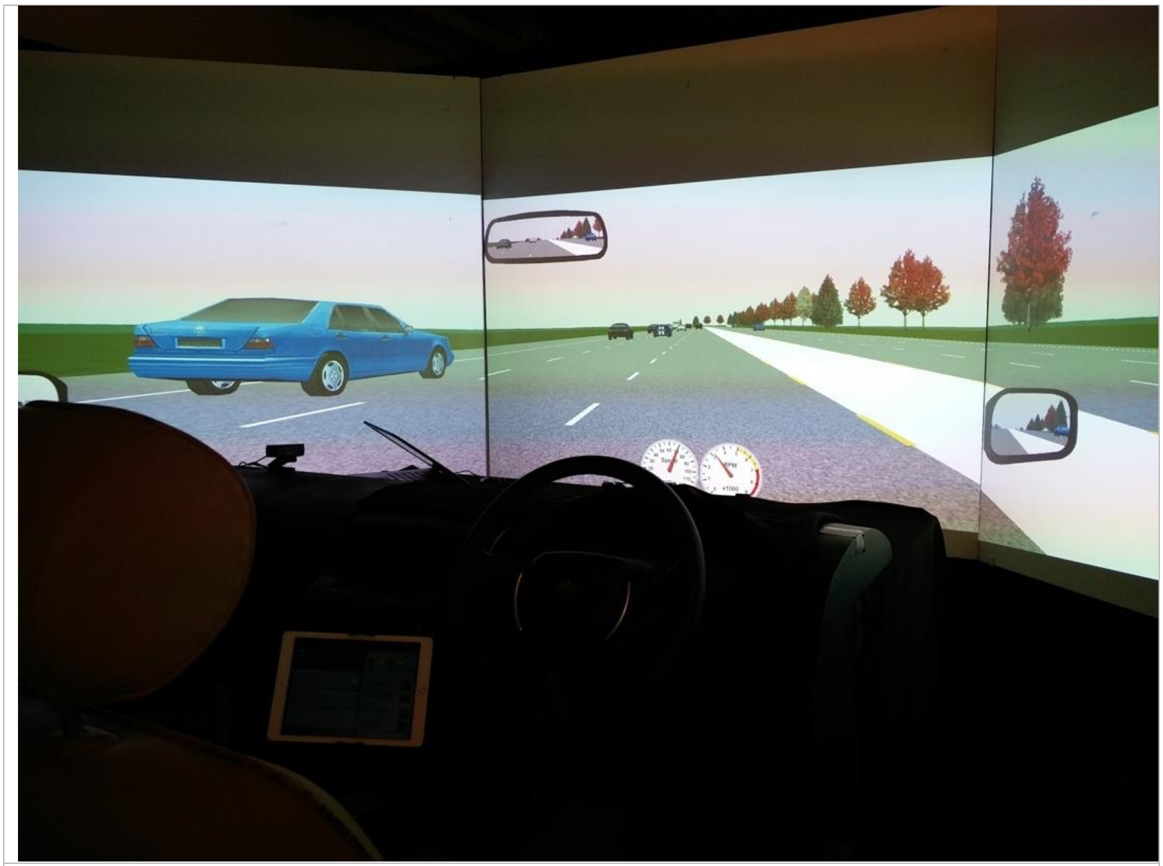


Figure 10: Photo of the driving simulator during an experiment.

3.3.3 The STISIM Driving Simulator

The STISIM driving simulator (STISIM, 2018) provides 135° with graphics projection serving as a test environment. The rig consists of an SUV seat and steering wheel with automatic transmission, see Figure 10. The cockpit included a tablet for multimedia use. On the right side of the cabin, a camera was placed to record the participant's posture and behaviour.

To communicate the transition from normal to automated driving, an audio message was played informing drivers that the system is taking over control when a certain point in the simulated environment is reached. When the ego vehicles get close to a predefined hazard for the scenario, an intermittent beep, based on NHTSA guidelines (Campbell et al., 2007), was played to instruct the driver to takeover. When the automated driving system is activated, the vehicle speed is set to 70mph and is placed to provide a seven seconds gap from the leading vehicle. A seven-second gap was chosen based on the study by (Gold et al., 2013a). The automated system provided lateral and longitudinal control with no overtaking manoeuvres, changing lanes or changing speed. This was done to reduce the number of independent variables and ensure drivers will engage in NDR tasks.

3.3.4 The justification for choosing the STISIM Driving Simulator

The driving simulator provided a safe approach to test reproducible driving scenarios to test driver’s ability to handle critical driving incidents when they are distracted or under time pressure to avoid a collision. Thus, a real driving approach was rejected because it will not provide enough safety for drivers and may put participants’ lives in danger. More details are furnished in section, section 3.2.1.

The STISIM driving simulator provides the necessary infrastructure to design and implement driving scenarios. The STISIM software uses a scripting language called the Scenario Definition Language (SDL) (STISIM, 2007). The SDL allows the design of the scenario, road, vehicles, speed limits and allows an extensive data collection of telemetry (e.g., throttle, braking pedal, steering angle, etc.) in addition to partial automation which was adequate for the experiment design.

One of the main challenges imposed by the STISIM software was the inability to change the angle of the steering wheel automatically through the STISIM software. To explain, when the automation starts, drivers may leave the steering wheel with an angle which is not changeable by the STISIM software. This posed a challenge when the vehicle switched from automation to at the transition because the vehicle would steer away from the hazard (vehicle ahead) putting the vehicle in an unknown state and therefore breaking the consistency of the experiment among participants. To solve that, an automated audio message was played stating 'Automation starts, please centre the steering wheel' when automation started to remind the participants to centre the steering wheel; thus, eliminating this problem. This was done during the training scenario and communicated orally with the participants to ensure they comprehended the issue. Then the aforementioned audio message was played at each transition from manual to automation during the main experiment.

3.3.5 Non-Driving Related Tasks

To understand how visuo-cognitive and cognitive distractions may affect physiological behaviour of participants, non-driving related tasks were selected based on two previous studies (Merat et al., 2012; Zeeb et al., 2015); namely email and twenty questions tasks. Also, participants in the control group were requested to pay attention to the road without engaging in any tasks. The sequence of those tasks was picked randomly for each participant to alleviate the order effect. The email task included reading an email on a tablet on the vehicle's dashboard then writing a reply. The emailer asked participants to pick a close friend and describe their perfect birthday party.

The Twenty Questions Task (TQT) was chosen because it causes a cognitive distraction and has been selected in similar studies (Gold et al., 2016; Merat et al., 2012). During the TQT, participants were asked to guess an animal by asking the researcher a maximum of 20 polar questions via a simulated hands-free phone call (Jamson et al., 2004). These two tasks were designed to ensure participants' engagement until the takeover request (similar to (Zeeb et al., 2015)) to maintain the same effect on mental workload among all participants. The duration of these tasks was seven minutes, and their order was randomised.

3.3.6 The justification for the Chosen NDR Tasks

According to Craye and Karray, (2015), driver distraction in manual driving could be divided into three categories that impose visual, cognitive and manual inattention -- they also reported that every secondary task (or in this study's case, NDR tasks) might include at least one of the aforementioned distractions. As discussed in section 2.4.2, visual and manual distractions are straightforward to detect using available technologies, see section 2.4.4 for more details. The main challenge was in detecting cognitive distraction because most indicators are indirect cues to the mental state and may be affected by other individualistic factors, see section 2.4.3 for more details.

Therefore, the study focused on the NDR tasks that could impose cognitive distraction. Among the surveyed NDR tasks, (see section 2.3.2), the study aimed at choosing distracting tasks that could stimulate cognitive distraction alone or visuo-cognitive distractions at the same time. These two tasks would be compared to a control group that performed no tasks to understand the following:

1. The effect of no tasks on physiological behaviour of drivers and compare it with their performance,
   - This was selected as the baseline of the study to compare others to those which were performed by the control group.
2. The effect of cognitive tasks on driver's physiological changes and their performance,
   - This was selected to understand what cognitive distraction could impose on the driver's performance and how it may influence the physiological changes among different participants. This was performed by the TQT group.
3. The effect of visuo-cognitive tasks on physiological behaviour and how they affected their performance,
   - This was selected to understand if combining visual and cognitive distractions could worsen the driver's performance and see if the physiological changes could be detected accordingly. This was performed by the email group.

Choosing the cognitive and visuo-cognitive tasks was based on the literature review conducted by the study which is summarised at section 2.3.2. The review showed that the Twenty Questions Task was heavily adopted as the main cognitive distraction task in several highly automated driving studies. Nevertheless, no studies investigated the TQT's effect on the driver's physiological behaviour in highly automated driving scenarios. Thus, it was selected as the cognitive task of this study.

While TQT is heavily used in driving studies, it's still criticised because it's not a naturalistic task that drivers tend to perform in their daily driving routines. In consequence, the choice of the visuo-cognitive task was gravitated to naturalistic tasks.

Zeeb, Buchner and Schrauf, (2016) accentuated visuo-cognitive naturalistic tasks in their study. They compared the email, video and news tasks to study their impact on the driver's performance in highly automated scenarios. Their results identified a gap between the objective and subjective measures of the difficulty of the email tasks. To explain further, the email tasks were the most difficult task reported by participants; however, it wasn't the most difficult one according to the objective measures of the driver's takeover quality. Such contradiction was worth the investigation to apprehend whether drivers' physiological changes could have an answer to such inconsistency. Thus, the email task was chosen as a naturalistic visuo-cognitive task for this experiment.

3.3.7 The Ethical Procedure

All participants were handed a 'Participant Information Sheet' that included the experiment procedure, what to expect, how data is collected, stored and manipulated and the importance of the experiment. Once they finished reading, participants were given time to ask any questions regarding the experiment and the information sheet. Finally, participants were asked to sign an 'Informed Consent' document to accept the data collection and processing procedure in addition to a health screening questionnaire before starting the study. They were also informed they could leave the experiment at any time. The study followed the university's ethical procedure and used a slightly modified version of the Loughborough University's Information Sheet and Informed Consent templates. The information sheet and the consent forms are attached in Appendix C.

3.3.8 Driving Scenarios

After finishing the ethical procedure, participants were informed orally of how the experiment was run and encouraged to ask questions. Before showing them the simulator, participants were asked of their familiarity with Android or iOS operating systems. Based on that answer, the multimedia tablet placed in the vehicle rig was chosen accordingly. This was done to alleviate any learning curve of handling the tablet during the experiment.

Participants were given 20 minutes of manual driving to get them familiar with the driving simulator. Then, they were given 10 minutes of highly automated driving practice before starting the main experiment. First, they were shown the rig and explained the features and limitations of the driving simulator. After 20 minutes of driving manually, participants were informed they could finish the training period whenever they felt fully accustomed to the driving simulator within a limit of two minutes. All participants chose to end their training within one minute of the request. The next training step involved the activation of the automated system.

Before the training phase starts, participants were required to put on an eye-tracking system (Tobii Pro Glasses) to track their eye movements and a heart rate monitor (Polar H7 chest strap). The eye tracking system captured the pupil diameter and saccade eye movements of the participants. The heart rate monitor was chosen for this study because it has a 99% accuracy in comparison to ECG devices (Wang et al., 2016). Though, it was difficult to set up for a few female participants because it interfered with their brassiere. To overcome that, those female participants were requested to wear a sports brassiere that had no metal support. A chaperone was present for the duration of the experiment. The chaperone was the same gender as the participant and helped place in the chest strap of the HR monitor. In addition, they were asked to put on a smartwatch with a heart rate monitor (Polar A360, (Polar, 2016)) as a second heart rate monitor in case the chest strap HR monitor failed. No data were collected during the training stage; fitting the data collection equipment was done to get participants used to their physical presence.

After the manual driving training ended, participants were given a chance to experience the automated driving once the equipment was all in place. Participants started a 10 km practice drive in a UK based three-lane motorway scenario.

At the second kilometre of the practice, automation starts. The automated system sets the speed at 70mph following a large SUV on the leftmost lane whilst setting a 500-foot distance from it. At kilometre eight, the TOR is automatically activated due to a broken-down vehicle in the current lane. The ego-vehicle starts decelerating immediately allowing seven seconds before a collision with the SUV assuming no driver's intervention. The scenario was defined in a manner which results in a collision with the broken-down vehicle if no action was taken by the driver allowing 7 seconds of reaction time as recommended by (Gold et al., 2013a). After finishing the takeover process, participants drove manually for two kilometres before the automation system is re-engaged. When participants were done, the experimenter asked them to relax for one minute before starting the main experiment. This was done to ensure the drivers' heart rate returns to a resting level.

The main experiment was approximately 30 minutes long. It included three takeovers and three slots for secondary tasks with no stop in between. The main scenario of this experiment was a repetition of the same practice scenario where participants started driving manually and then placed the vehicle on automated mode at the second kilometre. This was done by asking the drivers to place the ego-vehicle in a specific lane. Lane choice was alternated every time to cover the three lanes for every participant; this was done to reduce expectancy effects and to consider takeover behavioural changes in different lanes. After a minute, drivers were asked (via an audio message) to engage in one of the NDR tasks or pay attention to the road.

The secondary task engagement lasted for approximately 7 minutes. Then, an intermittent beep was sounded to a signal a takeover request. Participants were expected to stop the NDR task and engage immediately in handling the vehicle. This scenario was then repeated twice to cover two more takeovers and two more slots for the NDR tasks. The main difference among these phases are 1) the NDR task; participants are expected to do a different NDR task in each phase and 2) the lane; the vehicle changes lane right after the automation starts. This repeated measure approach allowed more results for each participant and provided data to study the learning effect of drivers and how it correlates with their physiological measures.

The driving scenario is explained more in details at Appendix A. It includes a visual explanation of the scenario with the key events in each step. In addition, it has the full source code of the SDL that defined the road, vehicles and the takeover process.

3.3.9 The justification for using the Repeated Measures Approach in the Driving Scenario

The author decided to adopt the repeated measure approach for the following reasons. First, it allowed more statistical power for a smaller number of participants. As stated in section 3.3, data collection was costly because of the specified dates in which the study could take place in addition to the difficulties in recruiting participants. Second, the repeated measure approach provided the necessary data to study the learning effect on the driver's performance and its potential effect on their physiological data. In few studies, a learning curve of driver's handling the takeover was observed; (Körber and Bengler, 2014; Larsson et al., 2014; Wright et al., 2016a). This phenomenon raises questions about whether drivers will maintain the same physiological behaviour consistently through the three takeovers or perhaps time may have an effect on those changes. Finally, the repeated measure design provided the possibility to study the order effect of the NDR tasks on the driver's performance.

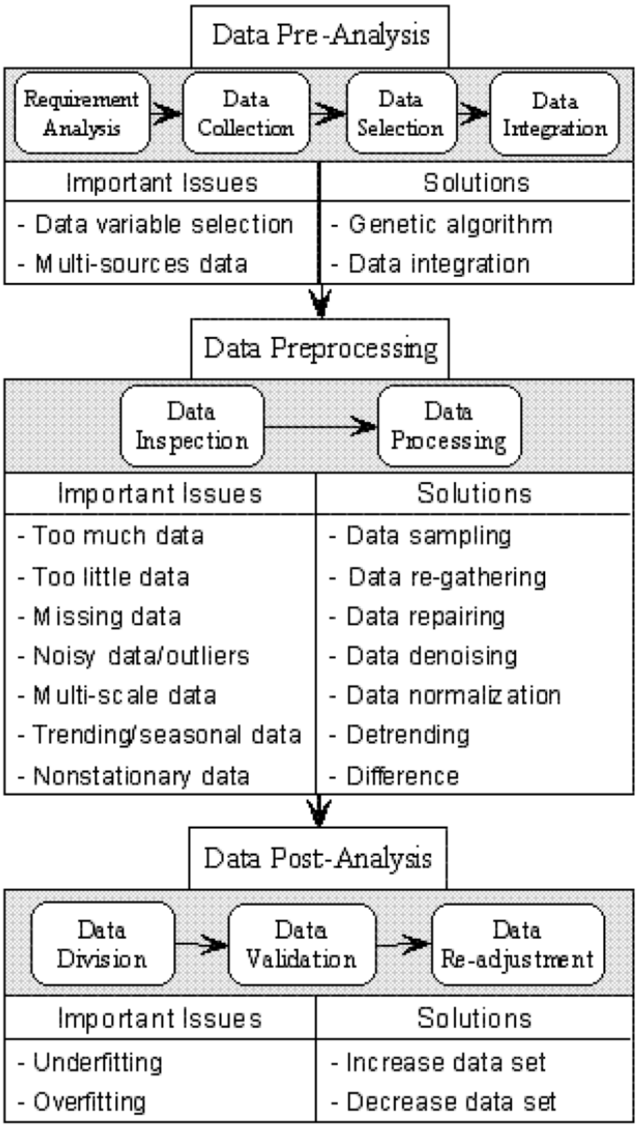


Figure 11: Data analysis scheme (Yu et al., 2006).

3.4 Data Analysis

3.4.1 Introduction

The quantitative data analysis method was chosen for this research to assess the correlation among the dependent and independent variables. As identified by Yu et al. (2006), data analysis could be separated into three main steps: data pre-analysis, data pre-processing and data post-analysis (see Figure 11). The study mixed the two approaches of Yu et al. (2006) and Kotsiantis et al. (2006) to develop a pre-processing framework adequate for the research problem. Yu et al. (2006) and Kotsiantis et al. (2006) studies were chosen because they were widely used in the literature. While Yu et al. (2006) created their framework for neural network modelling, the system could be easily extended to other machine learning techniques. The steps taken for the data analysis are explained in detail in the next few sections.

3.4.2 Data Pre-Analysis

Data pre-analysis is the step-in which data of interest is identified and captured (Yu et al., 2006). Yu's et al. (2006) framework consists of the following steps: requirement analysis, data collection, data selection and data integration. At the requirement analysis step, the problem targeted by the study should be identified clearly to collect a data requirement scheme. The scheme helps researchers to understand what information to collect, which data is required for each task, the data format, internal and external sources of data.

Once the requirement is set, the data collection step is performed. The importance of this step lies in the fact that data collection should avoid any biases or errors as it may resonate throughout the data analysis pipeline. Choosing an adequate data variable is essential to find the simplest facts representation based on the lowest number of variables. This ensures saving time in modelling and reduces the problem space (Stein, 1993, Kotsiantis et al., 2006).

Finally, data integration is important when several data sources are used to collect data. At this step, data is combined, merged, and synchronised from such sources. Based on the nature of data, it could be stored on structured or non-structured databases.

3.4.3 Data Pre-Processing

Data pre-processing is the step that data has to be transformed and conditioned to filter inaccurate and divergent data. This phase consists of two main steps: data inspection and data processing. Inspecting data means finding any issues with data quantity and quality. Generally, some data could be too small, too big, noisy or missing. Solutions to these issues include data regathering, sampling and linear regression in order (Yu et al., 2006).

At the data inspection step, data is analysed to find issues such as missing data points, trending data, or data size issues (Yu et al., 2006). In addition, data representation could be unsuitable for some statistical models; therefore, a normalisation step is required (Kotsiantis et al., 2006) at the data processing step. Moreover, missing data points, illegal values (for cardinal metadata for example) and out of range for some sensory data have to be identified and tackled (Kotsiantis et al., 2006).

Table 7: Summary of the dependent variables used in the experiment.

Measure	Source	Measurement Type	Range
Response Time	Driver Simulator	Seconds	0-7
PerSpeed	Processed Data of the Driving Simulator	Percentage Change	Starts at 0 which indicated no change in speed. Then goes up based on the observed changes in speed.
PerAngle	Processed Data of the Driving Simulator	Percentage Change	Starts at 0 which indicated no change in the angle of the vehicle. Then goes up based on the observed changes in the vehicle's heading angle.
Takeover Difficulty	Survey	Nominal	Difficult, Neutral and Easy
Distraction Level of each NDR Task	Survey	Ordinal	1-5 where 1 means extremely not distracting to 5 which means extremely distracting
Heart rate	Polar H7 Chest Strap	Beats Per Minute	0 – 1. That's because heart rate data was normalised.
Pupil Diameter (left and right eyes)	Tobii Glasses	Millimetre	0 - 1. That's because pupil diameters were normalised.
Saccade eye movement	Tobii Glasses	Magnitude of X,Y,Z movements of the eye.	Starts at 0 which indicated no eye movements detected. Then goes up based on the saccade eye movements. A high value indicated rapid eye movements.

3.4.4 Data Post-Analysis

At this phase, data is ready to be used for building up classification or regression models. Data is divided into mainly two to three sets. The training set is used to teach the model the trends of data. Then, the validation set is used to test and train the model. This step is optional and could be skipped. Finally, the test set is used to test the efficiency of the model. Typically, data is split into 70% training data set, 20%

validation set and 10% testing set (Yu et al., 2006). Other methods exist for small datasets such as N-Fold Cross Validation, see section 3.4.8.9.

In addition, feature selection is required to identify and remove any redundant features fed to the statistical model. Features are labelled relevant, irrelevant, or redundant. Relevant features are identified as the ones that have an influence on the output and cannot be identified by other features. In contrast, irrelevant features have no influence on the selected output. Redundant features provide information that has already been identified either explicitly or implicitly by another feature (Kotsiantis et al., 2006). A widely accepted approach for feature selection was proposed by Blum and Langley (1997) that uses a heuristic model for grouping and evaluating the relevance of each feature. More information is furnished in section 3.4.9.

Table 8: Summary of the independent variable used in the experiment.

Measure	**Source**	**Measurement Type**	**Range**
Type of Distraction	Research Choice	N/A	Email. TQT or control group

Data validation is the next step in this phase. Models require several training epochs to reach an acceptable training error. Adapting the model and retraining it is necessary to achieve an acceptable result (Yu et al., 2006). Models could suffer from two main issues, underfitting and overfitting. Overfitting happens when the model cannot generalise, i.e., the model performs well on training data but poorly perform on test data. Conversely, underfitting happens when the model is unable to fit well to the training data and is not capturing trends on the training data. This is also called over-generalisation. A perfect model has to have a balance between these two issues to ensure adequate generalisation to perform consistently (Yu et al., 2006, Van der Aalst et al., 2010).

Finally, based on model performance, a data readjustment step could be needed. When the model is underperforming, data must be readjusted to overcome the overfitting and underfitting problems. To resolve fitting issues, the model has to be retrained with different data preparation strategies such as changes in the data splitting, the model hyperparameters or learning iterations (Kotsiantis et al., 2006).

3.4.5 Problem Definition

During the driving simulator experiment, heart rate (HR) and pupil diameters (PD), the location of pupils of both eyes were collected before a takeover request (TOR) occurs. Then, Driver's response time and handling of the vehicle was recorded. PerSpeed and PerAngle are two performance measures that were introduced and evaluated in the study, see Chapter 4. PerSpeed is defined as the mean percentage change of vehicle speed whereas, PerAngle is the mean percentage change of the heading angle. The aforementioned measures were used as the variables assessing the takeover quality. Table 7 shows a summary of dependent variables collected in this experiment. The main independent variable in this experiment was the type of distraction. The study collected heart rate, pupil diameter and movement in addition to some surveyed data such as subjective level of distraction and their order of difficulty, more details are provided in Table 8.

Table 9: Example Data Collected During the Experiment

Takeover ID	Time (ms)	HR	Right eye				Left eye			
			Pupil diameter	Location X	Location Y	Location Z	Pupil diameter	Location X	Location Y	Location Z
1	0	74	0.01	0	0	0	0.015	0	0	0
1	1	74	0.012	0.2	0.1	0	0.016	0.22	0.12	0
1	2	74	0.016	0..4	0.4	0.2	0.017	0.44	0.41	0.19

The study investigated the predictability of the driver's performance measures in highly automated driving using physiological measures data. The study also identified the most relevant features that could be extracted from the collected dataset. An example of the dataset is illustrated in Table 9.

Table 9 shows a sample of the multivariate time series data of HR, PD and location of pupils of both right and left eyes of participants collected over time. Time series is a series of variables identifying the values of a variable over time and multivariate refers to the fact that the data includes multiple time series (Anderson, 2011). The dataset contains 99 takeover instances collected from 33 participants (53% females) between 20 and 30 years of age (M=25.8, SD=5.7). Data were collected up to 200 seconds before each takeover request.

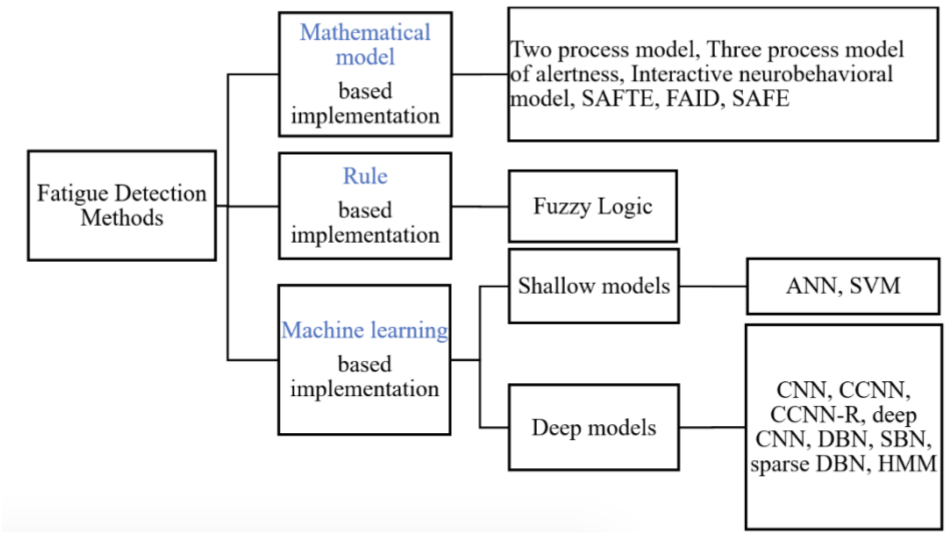


Figure 12: Different approaches to driver fatigue detection (Sikander and Anwar, 2018). The study used the same approach to categorise different detection algorithms of driver's inattention.

Each takeover had an ID assigned to it. HR was collected at 1Hz and PDs were collected at 60Hz. Out of six participants, 21 HR readings were missing out of an average of 2000 readings. To fill HR missing points, they were interpolated using a linear regressor to fill gaps between readings. For each takeover ID, there was another table that had a response time, PerSpeed and PerAngle values associated with it.

The dataset nature of being a multivariate time series imposed several challenges. To begin with, each takeover is represented by an average of approximately 30,000 readings of pupil diameter, eye movements and 200 readings of HR data. Using the raw data as an input to any data analysis approach (Machine learning for example) has major drawbacks such as sensitivity to noise in data, different lengths of time series and mismatch in correlating events happening at different times (Nanopoulos et al., 2001). To alleviate that, literature in time series analysis field suggested a feature-based approach for regression (Kléma et al., 2004) and classification (Bagnall et al., 2017; Nanopoulos et al., 2001). In addition, the literature produced other approaches such as MUSE (Schäfer and Leser, 2017a) to create a domain agnostic multivariate time series classifier by encapsulating features generation into their processes. The two approaches are explained in more detail below.

Feature-based approach converts time series into first and second order levels of features (Nanopoulos et al., 2001). Features are statistical values that aggregate information over a period of time and are not dependent on a specific time point which makes features less sensitive to noise (Fulcher and Jones, 2014; Nanopoulos et al., 2001). In addition, the aggregation of those time points reduces the machine learning models input and therefore reduces learning time considerably. Then, several machine learning algorithms are applied to perform classification using the generated features. More details will be discussed in 3.4.6 section.

From driver inattention studies' perspective, the literature review identified three different approaches for driver inattention detection: mathematical, rule and machine learning-based approaches (Sikander and Anwar, 2018). While Sikander and Anwar's, (2018) literature review covered only fatigue studies, their categorisation approach could be extended to driver inattention studies. For example, driver inattention methods this study surveyed included mathematical-based approaches (Schmitt et al., 2018), in addition to rule-based approaches (Sigari et al., 2013) and machine learning-based approaches (Jafarnejad et al., 2018). In summary, the model displayed in Figure 12 summarises all approaches used in driver inattention studies.

There are several advantages to the machine learning approach in comparison to the mathematical and rule-based models. For example, the machine learning approach automates the knowledge acquisition from the collected data (Smola and Vishwanathan, 2014). Conversely, mathematical models require a heavy domain-expert knowledge (Mallis et al., 2004) which make them difficult to produce and evaluate since the field of highly automated driving is relatively new (Campbell et al., 2018). Similarly, rule-based models using Fuzzy Logic are dependent on function approximation which is practically another form of interpolation; hence, they are only suitable for problems that have an adequate mathematical description (Reus, 1994). Accordingly, the mathematical and rule-based approaches were considered unsuitable for assessing the predictability of driver's takeover performance.

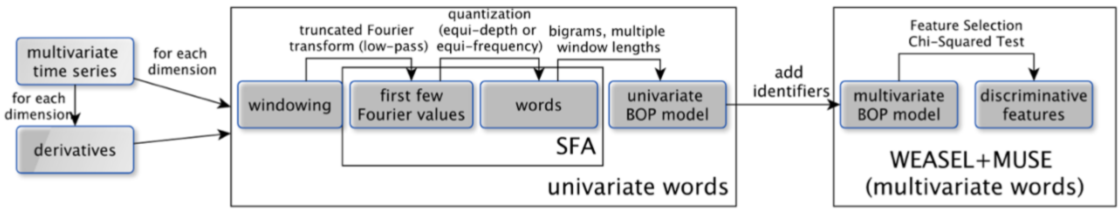


Figure 13: "WEASEL+MUSE Pipeline: Feature extraction, univariate Bag-of-Patterns (BOP) models and WEASEL+MUSE", (Schäfer and Leser, 2017a)

Contrarily, machine learning-based methods provide more flexibility in developing predictive models (Smola and Vishwanathan, 2014) which made them applicable for several driver inattention studies (Dasgupta et al., 2013; Jafarnejad et al., 2018; Le et al., 2018; Li et al., 2011; Liang et al., 2007; Solovey et al., 2014). Several machine learning-based studies achieved high accuracies in driver inattention detection, see section 2.4. Subsequently, the study decided to choose the machine learning approach.

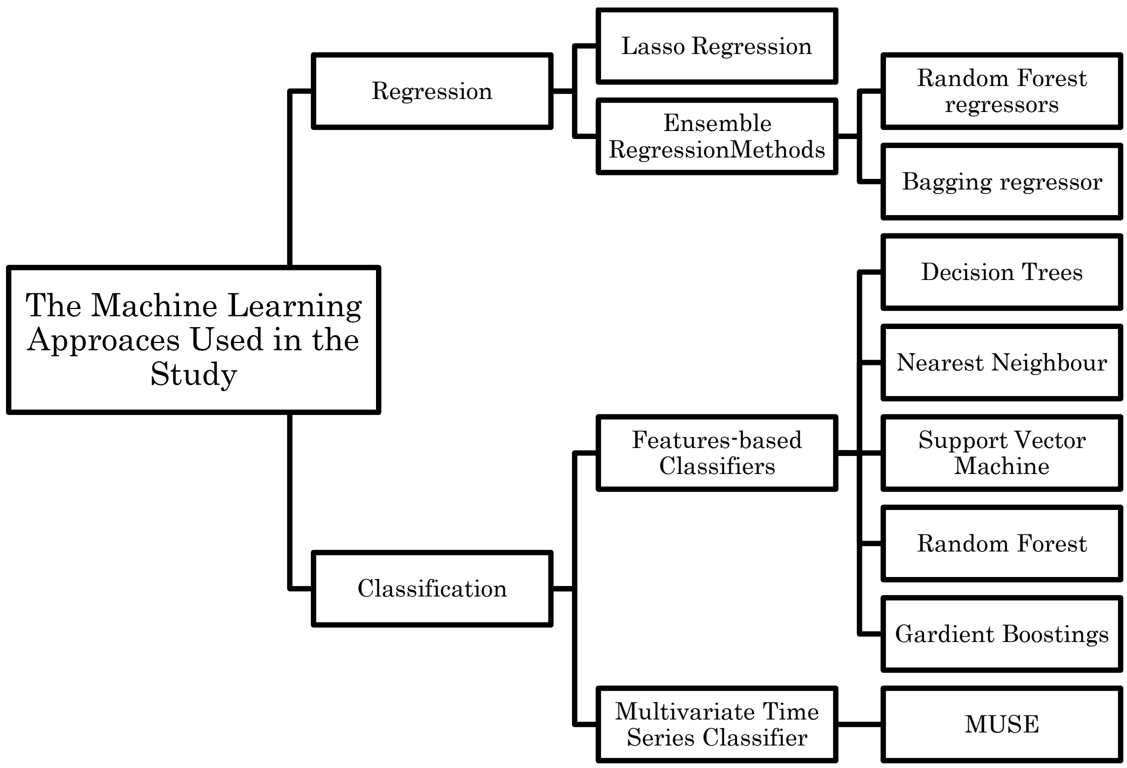


Figure 14: An Illustration of different machine learning models used in the study.

Contrarily, there's a family of multivariate time series classifiers that encapsulates feature learning and classification (Bagnall et al., 2017). For example, MUSE is a state-of-the-art multivariate time series classifier. It performs its own feature extraction and learning using bag of patterns approach to convert time-series frequencies into words using Symbolic Fourier Approximation (Schäfer and Högqvist, 2012), then filters them based on their information gain. At the end of the features filtering, the algorithm uses linear classifiers as a final classification layer. The full pipeline is displayed in Figure 13.

The individuality of recordings of HR and PD imposed another challenge on providing a driver-independent predictor. Recruited participants had different resting heart rate and different mean pupil diameter. To alleviate the individuality of the data in order to make them comparable, data normalisation was required for better comparison among participants (Patro and Sahu, 2015). More details are provided in 3.4.8.1.

3.4.6 The Statistical Approach

The study used repeated measures ANOVA and linear mixed models (LMM) to study the difference among group means and study the correlation between the dependent variables and their covariates. Repeated Measures ANOVA was used to test the statistical significance between NDR tasks and their corresponding mean heart rate and pupil diameters of participants. In addition, it was used to test the statistical significance among different PerSpeed and PerAngle groups. PerSpeed and PerAngle groups were determined using K-Means Clustering algorithm.

Linear Mixed Models were used to assess the correlation between response time and their corresponding mean heart rate and pupil dimeters while considering the repeated measurability nature of the data. In addition, PerSpeed and PerAngle were tested using LMM to assess a correlation with the mean heart rate and pupil diameters of participants.

The human factors field has adopted ANOVA approaches to identify the statistical significance among groups, see Chapter 2 for the literature review. Additionally, Linear Mixed Models were used in similar studies, e.g., (Cummings and Guerlain, 2007) to identify the correlations among continuous variables.

3.4.7 The Machine Learning Approach

3.4.7.1 Introduction

Using machine learning, the study used two methods to evaluate the predictability of drivers' performance measures, see Figure 14. The target variables were Response Time, PerSpeed and PerAngle. To predict those variables, two approaches were used. The first approach was regression-based to predict the actual values of the performance measures. The second approach was classification-based where performance measures were categorised into discrete values. In the following sections, regression and classification methods used in this study are explained. The justifications of each method used are briefly discussed.

In the regression approach, two methods were used:

3.4.7.2 Linear Regression

Linear regression is a statistical technique to assess the linear relationship between dependent variables and one or more independent variables (Neter et al., 1996). It has been previously used in assessing fatigue (Li-Wei Ko et al., 2015), visual and cognitive distraction (Li and Busso, 2015).

3.4.7.3 The justification for Using Lasso Linear Regression

Lasso algorithm is a linear regression model optimised for prediction. The algorithm uses regularised coefficients which enables it to generalise for new datasets and therefore, build complex models while avoiding overfitting (Tibshiranit, 1996) which makes it more advantageous over regular regression methods (Hansheng et al., 2007).

3.4.7.4 Ensemble Regression

The ensemble regression models are meta-estimators that combine multiple regressors on a patch of sequential or random subsets of the training set (Tin Kam Ho, 1998). Their individual predictions are then aggregated to form a final result. The randomisation added into the construction of those methods reduces the variance in comparison to other estimators such as Gini tree methods which alleviate overfitting (Breiman, 1996).

3.4.7.5 The justification for Using Ensemble Regression

In the field of driver inattention, several studies proved that ensemble regression is adequate for classifying driver mental workload (Le et al., 2018) in addition to fatigue (Hu and Min, 2018). The ensemble regression learning models are known to enhance learning performance by combining multiple models (Berk, 2006). Consequently, they provide better results than regular regression models due to their ability to adapt to highly dimensional datasets (Wang et al., 2010). Therefore, ensemble regression was used along with linear regression to compare their results and explore the dimensionality of the collected dataset.

3.4.7.6 Feature-Based Machine Learning Classifiers

The study used the spiral system development approach, as explained in section 3.2 and justified in section 3.2.3. The accuracies of the developed regression and ensemble regression methods in the study were inadequate for some of the performance measures, refer to Chapter 4 for more details. Thus, the author decided to investigate the classification approach to build predictive models for the performance measures. The study investigated two methods: feature-based and multivariate timeseries classification methods. In the following paragraph, a feature-based method is explained and justified. Section 3.4.7.8 discusses time-series based classifiers, and the justification for their use is furnished.

Classification is the process of mapping input data into a class in a predefined set of classes (Smola and Vishwanathan, 2014). Classification requires a set of predefined classes outlining the output of the model for every corresponding input data. In addition, the training of the model requires a set of well-chosen features (input data) that provides a high information gain to the classification model (Nanopoulos et al., 2001). Choosing a set of features is crucial to the robustness of those classifiers (Nanopoulos et al., 2001). Consequently, data preparation is essential in creating robust classifiers in addition to choosing the classifiers themselves. This section is concerned with the choice of the classification model. For more information about data preparation, please check section 3.4.8.

The study used five classification algorithms; namely, Decision Trees using Gini algorithm (Quinlan, 1986), Random Forest (Breiman, 2001), Gradient Boosting classifier (Freidman, 2008), Support Vector Machines (SVM) (Suykens and Vandewalle, 1999), and K-Nearest Neighbour (KNN) (Cover and Hart, 1967) were used.

3.4.7.7 The justification for Using Different Machine Learning Classifiers

According to the 'No Free Lunch Theorem', there are no machine learning classifiers that perform best for every problem; hence, comparing several classifiers is a necessity to identify the most suitable one for the study's problem (Wolpert, 2002). The practice of comparing multiple classifiers is widely adopted in several driver inattention studies (Darzi et al., 2018; Jafarnejad et al., 2018; Le et al., 2018; Solovey et al., 2014). Hence, the study chose a collection of classifiers to test.

Picking the five classifiers was based on their reported performances in the literature. Decision Tree classifier was used because of their ability to provide a human readable solution and for their easiness and training speed (Quinlan, 1986). SVM classifier was used because of its ability to handle high dimensional data (Bennett and Campbell, 2000; Dong et al., 2015; Smola and Schölkopf, 2004). Default hyperparameters based on Scikit-learn models, a Python library (Pedregosa et al.,

2011) were used to obtain unbiased results and avoid overfitting. To benchmark the classifiers results, Gaussian Naïve Bayes (Zhang, 2004) and random classifier algorithms were used as a baseline.

3.4.7.8 Multivariate Time Series Classifiers

In the literature of time series classification, generic approaches (dictionary-based as referred to in some papers (Bagnall et al., 2017)) were developed and tested on several timeseries. The generic classifiers encapsulate feature learning, selection and training to create domain-independent time-series classifiers (Schäfer, 2015; Schäfer and Leser, 2017a, 2017b). While univariate time series classifiers research has progressed significantly in the past decade, (Bagnall et al., 2017), research in the field of Multivariate Time Series classifiers is limited according to the review of Schäfer and Leser, (2017b) who proposed a novel MTS classifier called MUSE.

3.4.7.9 Justifications for Using Multivariate Time Series Classifiers

The study chose MUSE for mainly two reasons: 1) it performed better than all other algorithms in domain agnostic time series classifiers (Schäfer and Leser, 2017a) and 2) it provides a comparison between the feature-based approach and domain-agnostic approach. The results of MUSE were later compared with the suggested feature-based solution in section 3.4.7.6. The authors of MUSE reported that their algorithm outperformed all its counterparts (Schäfer and Leser, 2017a) which is verified on a public code repository on Github (Patrick Schäfer, 2017). In this study, MUSE is used and compared against the feature-based approach on the collected dataset.

3.4.7.10 Summary of the Machine Learning Approach

The study adopted the machine learning approach to build predictive models for the driver's performance measures. The study used two methods: regression and classification. Regression methods included simple regression models such as Lasso and ensemble regressors such as random forest and bagging. The second method used was classification. It was adopted because regression models underperformed in predicting some of the driver performance measures. Feature-based classifiers were used: decision trees, KNN, SVM, RF and GB. Finally, a multivariate time series classifier called MUSE was used to compare its accuracy to the feature-based approach.

3.4.8 Data Preparation

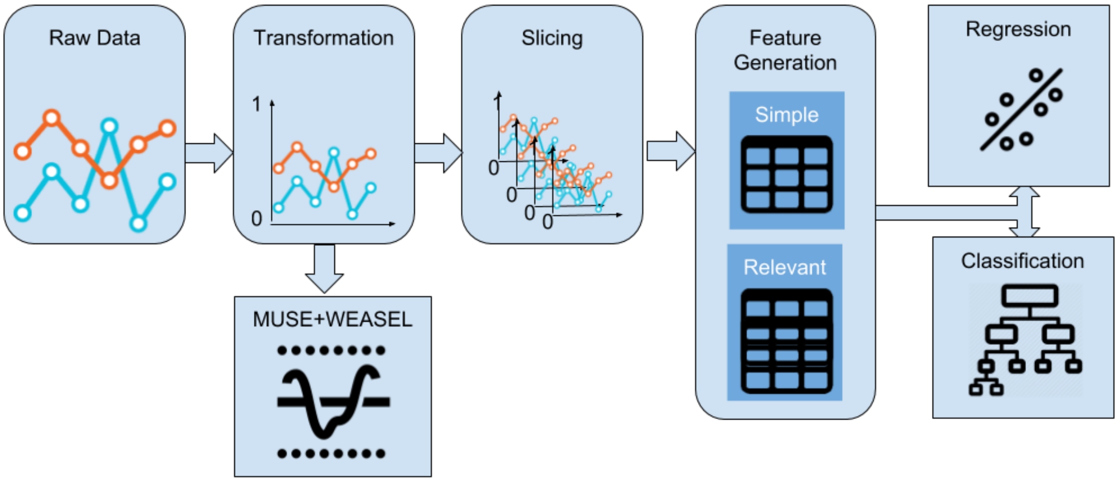


Figure 15: An illustration of the data preparation process. Raw data is transformed by normalising the data to fit in the range from 0 to 1. Then the data is sliced into window sizes. Different features are extracted using simple and relevant approaches. Finally, the generated features are passed to either regression or classification algorithms.

Data preparation is a crucial step to create robust predictive models (Bramer, 2016; Dalgleish et al., 2007). The study put a significant effort in the data preparation by creating a general framework motivated by previous research (Fulcher and Jones, 2014; Kléma et al., 2004). The framework is illustrated in Figure 15 and explained in detail in the following sections. The framework consists of four steps, normalisation, transformation, window slicing and feature generation. In the following sections, each step is explained and justified.

3.4.8.1 Normalisation

Normalisation is the "adjustment of the values of an attribute generally to make them fall in a specified range such as 0 to 1", (Bramer, 2016). Pupil diameter (PD) and heart rate (HR) data were normalised according to Equation 1 where V is a window of the collected data ending at each takeover request. A window is a number of recordings that span over a period of time, e.g., a window size of 30 seconds includes 30 recordings of HR data over a period of 30 seconds at a sampling rate of 1Hz or 300 recordings of pupil diameter at 10 Hz. Normalisation was performed to alleviate the individuality of the collected data by putting minimum HR or PD value at 0.0 and a maximum value at 1.0 per each participant for the entire duration of the experiment.

$$\alpha = \frac{1}{n} * \sum_{i=k}^{n} \frac{V_i - \min(V)}{\max(V) - \min(V)}$$

Equation 1: Normalisation of heart rate and pupil diameters data were performed using the method, where V is a vector containing readings from k is time0 and n is window size.

3.4.8.2 The justification for Using Normalisation

The normalisation process facilitates the comparison among different participants. For example, different participants have different resting HR based on different variables (Karvonen and Vuorimaa, 1988). The recruited participants had a resting heart rate ranging from 39 to 75 beats per minute. Having a resting heart rate of 39 was alarming to the experimenter since it could either be a sign of fitness (Plowman and Smith, 2007) or a serious medical condition (Baliga and Eagle, 2008). When the participant was asked about their resting heart rate, they informed the experimenter they're a tri-athlete, and they're aware their resting heart rate is around 39. That wasn't the only incident during the experiment which is understandable since they were recruited at Loughborough University campus where the majority of students are involved in sports (Loughborough University, 2018). The reported range in the resting heart rate among participants makes it difficult to compare those values to each other. To mitigate, that, normalisation was adopted in the data preparation pipeline.

The study picked the simple min/max normalisation approach even though there are other popular normalisation techniques in the machine learning field (Zheng and Casari, 2018). Choosing a simple normalisation technique makes it easier to analyse and read the reported ANOVA results and could drive the adoption of normalisation in the field of Human Factors.

3.4.8.3 Sliding Window

The time series data are a series of values over time, meaning that every attribute has many values observed during a time period (Zheng and Casari, 2018). The format of time series is incompatible with the supervised learning methods that expect one value attribute (Nanopoulos et al., 2001). In order to fit time series data into supervised learning problems format, the values of a specific attribute are aggregated using some statistical formulas, e.g., sum, min, max and so on (Anderson, 2011). One of the challenges in this process is choosing the number of values to apply those formulas to – the number of values fed to the formula is called the 'window size'. Thus, the dataset was sliced into different window sizes. The window size ranged from 2 to 200 seconds.

3.4.8.4 The justification for Using Window Slicing

The window size was found to be crucial to the robustness of time series classification algorithms (Liang et al., 2007; Solovey et al., 2014; Wijsman et al., 2011). Thus, testing different window sizes was essential to identify their effect on the effectiveness of the reported models. The window sizes are defined as the time taken to collect data before a TOR. Previous studies indicated that the length of window size had a strong correlation with prediction accuracy (Solovey et al., 2014). Therefore, trying out different window sizes was important to explore this correlation in our dataset. However, overlapping windows did not add any performance gain (Solovey et al., 2014); therefore, the overlapping was not investigated in this study.

3.4.8.5 Feature Generation

A feature is defined as "a numeric representation of an aspect of raw data", (Zheng and Casari, 2018). A feature is generated by processing a window of a variable to create a single value, e.g., min or max (Anderson, 2011). Generating features in this study was the most crucial step to the robustness of the ML algorithms. Many studies explored suitable features in manual driving studies (Wijsman et al., 2011); however, highly automated driving studies are scarce, and the area hasn't been explored fully. Therefore, this study examined two methods for feature generation.

1) **Simple features generation:** The min, max, standard deviation, variance, length, sum and median were calculated for HR and PD time series which generates a total of 18 features for three time-series. Then, pupil locations (X, Y and Z) were later processed to calculate the magnitude of the eye movement vector. Several studies suggested that saccadic eye movement as a prominent feature for identifying mental workload (May et al., 1990; Tokuda, 2010). To calculate such feature, pupil location vector v_i represents the magnitude of the eye movement at time i using Equation 2.

Equation 2: the magnitude of eye movement formula.

$$\mathrm{v_i} = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2}$$

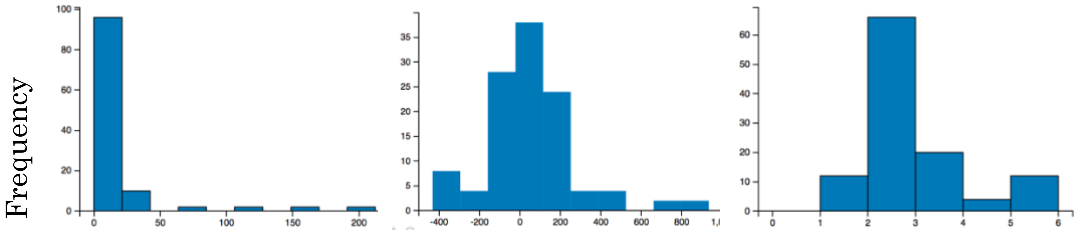
where v_i is the value of the saccade eye movements vector at time i, then the x, y and z are the coordinates of the pupil's location

2) **Relevant features generation:** A total of 600 features in temporal and frequency domains in addition to shapelets for the three time-series in the dataset were

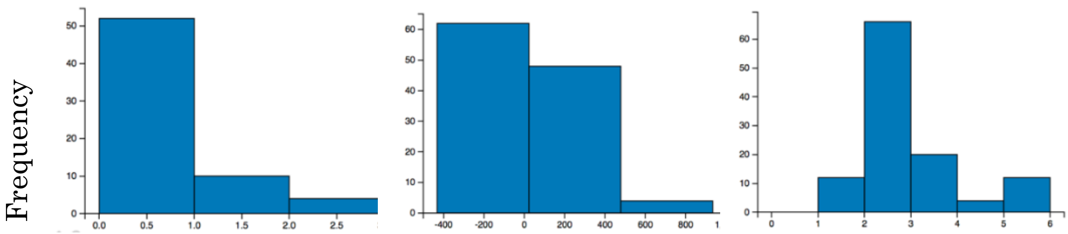
generated in this experiment. Such a large number of features requires filtering to find the most relevant features. Therefore, the study compared a few feature-selection algorithms to find the most suitable features subset to build the best classifier.

3.4.8.6 The justification for Using Two Feature Generation Methods

The simple feature generation method uses simple statistical functions. The intelligibility of the simple features aims at creating simple classifiers that could learn from the descriptive statistics of the input signals. Thus, providing an easy explanation to the researchers on how to interpret the changes in the collected variables over time.



PerSpeed Values | PerAngle Values | Response Time Values



Three bins representing categorised PerSpeed values | Three bins representing categorised PerAngle values | Six bins representing categorised Response values

PerSpeed | PerAngle | Response Time

Figure 16: Histogram of PerSpeed, PerAngle and Response Time values are plotted in the first row. The second row has the histogram of the same variable after being clustered. PerSpeed and PerAngle were clustered into three classes while response time was clustered into six classes.

Contrarily, the relevant features approach is a comprehensive feature engineering approach. The introduced approach is tailored to search for a set of features that could provide near-optimal classification accuracy. The process of relevant features searches allowed the researchers to explore and identify new features on shapelet, time or frequency domains.

3.4.8.7 Discretisation

The discretisation is "the conversion of a continuous attribute to one with a discrete set of values, i.e. a categorical attribute", (Bramer, 2016). It is used to convert continuous variables (e.g., Response Time) to categorical variables in order to reduce the number of values a continuous variable has by grouping them into a number of classes (Bramer, 2016). For example, continuous values of PerSpeed, PerAngle and Response Time were used in regression algorithms. Due to the unsatisfactory results of the regression methods, the author decided to explore classification algorithms. In order to fit PerSpeed, PerAngle into classification algorithms, their values were discretised into three categories using the K-Means algorithm.

K-Means clustering algorithm is a widely used supervised clustering algorithm that partitions a set of values into a *k* number of clusters where each observation falls into its nearest mean (Bramer, 2016).

3.4.8.8 The justification for Applying Discretisation

Classification algorithms typically take a set of classes rather than continuous variables (Bramer, 2016). Discretisation was applied to fit the performance measures variables to the classification algorithms. Choosing three categories was motivated by previous studies that split drivers into low, medium and high-risk participants in highly automated driving studies (Zeeb et al., 2015). With the three classes split, predicting the values of PerSpeed and PerAngle is still meaningful to identify the risky state of the driver at the takeover whether it could be low, medium or high risk when handling the vehicle. Similarly, response time was divided into six categories by rounding values to the closest integer, (e.g., $3.2 \approx 3$, $4.9 \approx 5$). Histograms of the performance measures are plotted before and after clustering on Figure 16.

Originally, there are two groups of discretisation algorithms: supervised and unsupervised (Kotsiantis and Kanellopoulos, 2006). Unsupervised approaches are used when the number of classes is unknown and typically figure out their own set of classes in which the data is split into. Conversely, supervised discretisation algorithms are used when the number of classes is known. The study used the supervised approach because it adopted the three class split of Drivers' Performance as discussed earlier from Zeeb, Buchner and Schrauf, (2015).

The K-Means discretisation algorithm was chosen because of its simplicity (Jain, 2010) and superiority to other algorithms (Finley and Joachims, 2008). In addition, its discretisation approach maximises the information gain and hence, eases the process of training (Kotsiantis and Kanellopoulos, 2006; Liu et al., 2009).

3.4.8.9 Cross-Validation

After pre-processing, the dataset was split into a training and test sets. Data splitting is the process of splitting the dataset into train and test sets. The train set is used to train the machine learning, and the test set is used to test the accuracy of the training model (Bramer, 2016). The study used the N-Fold Cross Validation which is also known as the 'leave-one-out' method. The definition of Leave-One-Out or "N-fold cross-validation is an extreme case of k-fold cross-validation, often known as 'leave-one-out' cross-validation or jack-knifing, where the dataset is divided into as many parts as there are instances, each instance effectively forming a test set of one.", (Bramer, 2016: 83).

3.4.8.10 The justification for Using N-Fold Cross-Validation

The dataset of the study was challenging due to its small size and unbalanced output. To counteract that, the study chose the N-Fold cross-validation. The main benefit of the N-Fold method is its suitability for small datasets since it allows for the maximum amount of data used to train the model (Bramer, 2016). In addition, Cawley and Talbot, (2004: 1467) stated that "Leave-one-out cross-validation has been shown to give an almost unbiased estimator of the generalisation properties of statistical models, and therefore provides a sensible criterion for model selection and comparison". The N-Fold technique hasn't been widely adopted in the literature due to its high computational cost with large datasets which is missing in this study because the dataset was small. Accordingly, the N-Fold technique was the most suitable cross-validation technique to be used.

3.4.9 Features Selection

3.4.9.1 Introduction

In similar studies, feature subset selection (FSS) is a well-known technique to pre-process the data for classification or regression (Kudo and Sklansky, 2000; Yang and Honavar, 1998a). FSS is defined as the process of filtering large features set into a smaller subset to maximise the prediction performance while maintaining efficient and fast processing of the data and predictor, e.g., classifier or regressor (Guyon and Elisseeff, 2003).

3.4.9.2 Feature Selection Method

In this study, TSFresh, a Python library was used to generate a total of 794 features on both time and frequency domains (Christ et al., 2018). Such a large number of features is expensive to process and may include strong collinearity among each other. The collinearity of features causes several classifiers and regressors to overfit or underperform (O'Brien, 2007). To mitigate that, a feature selection process was required to:

1. Filter out the colinear features to avoid any biases in our predictors,
2. Rank features according to their relevance,
3. Select a subset of the features that maximise the predictability of the predictor.

1) Filtering

The FRESH algorithm is a feature extraction algorithm that uses the Benjamini-Yektieli procedure to identify the relevance of features and filter out the irrelevant ones (Christ et al., 2018). When FRESH was run on the extracted features, a total of approximately 420 features (out of 794 that were generated) were selected. To reduce the number of features and maintain decent predictability of proposed predictors, colinear features had to be removed. There are several methods to assess the collinearity of features; Variance Inflation Factor (VIF) (Marquaridt, 1970) is known as a stable statistical measure and has been widely adopted in the literature

Variance Inflation Factor (VIF) is a method used to quantify the variance inflation caused by multicollinearity in the dataset variables. When two variables are colinear, their variance is subsequently inflated. VIF is calculated by estimating the ratio between the variance of all of the model's betas and the variance of a single beta of the model when calculated alone (Marquaridt, 1970). VIF, using a threshold of 5.0, was performed on a list of 420 features to identify the collinear variables. By keeping only one feature of all colinear groups, the number of features dropped to 120 features.

2) Ranking

To explore the importance of each feature, the study introduced a ranking approach to select the top features. For example, the top 50 features were chosen based on

Random Forest variable importance (Louppe et al., 2013). The 50 features were used as an input to the regression algorithms. However, the problem was more challenging in classification methods since the output classes were unbalanced. By picking top k features, generated models may be highly biased by picking the dominant output classes and as a result, reduce their precision and accuracy. Consequently, this study picked a selection of the most relevant features per class.

To explain the selection process in detail, features were ranked in multiple tables — each table corresponding to one output class based on their Random Forest variable importance (Louppe et al., 2013). To perform the relevancy test, a one-vs-rest approach was used to identify the top 10 features that provided the highest significance in prediction in a Random Forest classifier. This was done for each class in the output data. So, the number of tables produced were k tables, where k is the number of classes of the predictor. Then a combination of those tables could potentially produce an optimal feature set. To find the optimal combinations of features to feed into a predictor, an efficient search algorithm was required to find an optimal or near-optimal combination of features. A state space consisting of $10 \times k$ features makes it very difficult for a greedy algorithm to find the optimal combination of features subset. So, a search algorithm was required.

3) Selection

The study explored several search techniques. Researchers in the past few years in the ML field have adopted genetic algorithms as an efficient optimiser for optimising hyperparameters of machine learning (e.g., da Silva et al., 2018; Yang and Honavar, 1998). Genetic algorithms are evolutionary based optimiser that mimics natural selection processes (such as breeding, crossover and mutations) to find solutions for problems with large state spaces (Sastry et al., 2005). Genetic algorithms are commonly used to solve problems with large state spaces because they offer a fast approach to reach a near-optimal solution (Sastry et al., 2005). A drawback to this approach is their tendency to find a solution that overfits training sets and get stuck in a local optimum (Srinivas and Patnaik, 1994). To alleviate this, several precautions were performed.

First, Random Forest predictors (e.g., classifier and regressor) were used as an evaluation function because they don't tend to overfit and for their fast training speed (Breiman, 2001). Then, multiple initial runs were performed on a standard genetic algorithm to perform sensitivity analysis based on recommendations of Srinivas et al., (2014). Such analysis helped identify the suitable number of individuals per generations in addition to mutation and crossover probabilities. After the sensitivity test, a standard genetic algorithm was run for an initial 100 individuals and a total generation of 50. Mutations' probability was chosen to be 5% of the population and crossover was assigned a probability of 70%.

3.4.9.3 The justification for Applying Features Selection Method

According to Zheng and Casari, (2018: 38), "Feature selection techniques prune away non-useful features in order to reduce the complexity of the resulting model. The end goal is an efficient model that is quicker to compute, with little or no degradation in predictive accuracy". Since a large set of features is generated in the 'relevant features generation' pipeline, see section 3.4.8.5, a pruning technique was required to filter out the irrelevant features and identify a small subset of features that maximise the information gain of the classifiers while minimising the number of features fed to the model.

The FRESH algorithm was used to filter out irrelevant features at the start of the study's feature selection process. The FRESH algorithm was chosen because it's designed to facilitate the knowledge-domain acquisition process (Christ et al., 2018) which is missing in the field. The TSFresh algorithm achieved a robust accuracy in comparison to other filtering algorithms such as Boruta and LDA filtering algorithms (Christ et al., 2016) and have wide acceptance in industrial projects (Christ et al., 2018).

However, the filtering process of FRESH doesn't recognise the collinearity of filtered features as observed from their studies (Christ et al., 2018, 2016). Since passing colinear features to some classifiers bias their predictability (Nicodemus and Malley, 2009), it was important to filter out the most colinear features. Therefore, VIF was used to estimate collinearity among the filtered features. The VIF algorithm was chosen because it's "has been widely applied in the scientific literature to diagnose the existence of collinearity", (Salmerón Gómez et al., 2016, p.1).

Ranking up the variables was necessary to identify the features that provided the highest information gain to the classifiers. Knowing the features that provide higher predictability develop the acquisition of domain knowledge (Christ et al., 2018). Random Forest (RF) variable importance (Louppe et al., 2013) was used to rank up the features for a few reasons. "RF can handle huge numbers of variables easily. A global relative variable importance measure can be derived as a by-product from the Gini-index used in the forest construction with no extra computation involved", (Guyon et al., 2006, p.302).

To select a subset of features, genetic algorithms were used. There are several approaches in the literature that reported better performances in finding features subsets such as Floating Search Methods and Simulated Annealing (Guyon and Elisseeff, 2006). However, they are computationally expensive to run on a large feature set. Thus, genetic algorithms were the best choice because they can outperform other methods in problems with large state spaces, i.e., when the number of features is over 50 (Kudo and Sklansky, 2000), especially when the study's large state space.

3.5 Delimitations of the Experiment Design

The literature review in Chapter 2 has identified nine variables affecting driver's response in highly automated driving. The nine variables are summarised in the following paragraph for a quick reference, see section 2.3 for more details.

1. **Time budget:** the time given to drivers to respond to a hazard before a potential crash occurs has a strong effect on the driver's response.
2. **Driver distraction:** drivers' attention may be spent to an NDR task during the automation phase which in turn eradicates their situational awareness. This affects the driver's response when they are prompted to perform a takeover.
3. **Fatigue and stress:** Some studies reported a correlation between an increase in fatigue and the duration of the automation. Fatigued drivers are known to respond ineffectively in comparison to alerted drivers. Fatigue and stress could separate into two different categories since stress could be a pre-existing condition before driving. The work on these two variables are limited in the literature so the study decided to merge them together.
4. **Traffic situation:** traffic density affects the driver's response time because it creates a complex driving environment where drivers tend to spend a longer time to regain situational awareness.
5. **Road condition:** road conditions are concerned with the variables that are affecting road conditions such as fog limits drivers' sight which imposes a longer time to restore situational awareness and thus affect the driver's performance when they have a limited time budget.
6. **Driver background measures - Age:** drivers' age affects their judgment and response time.

7. **Driver background measures - Driving Experience:** experience with automated systems affect their complacency, trust in automation and response quality.

8. **Driver background measures - Individual differences:** drivers have different driving styles, and this is another factor affecting drivers' responses to critical hazards.

9. **Human Machine Interface:** the design of the communication channels between the automated systems and drivers played a major role in the distribution of driver's attention before the transition phase.

Several studies explored the effect of time budget on the driver's performance (Eriksson and Stanton, 2017; Gold et al., 2016, 2013a; Körber et al., 2015b; Zeeb et al., 2016, 2015). Recommendations in several studies suggested that 7 seconds is an adequate time for drivers to restore situational awareness and respond safely to critical hazards in typical driving scenarios. Based on that, the study decided not to vary the time budget variable. Moreover, the study aimed at varying the driver's distraction types in order to understand their effect on driver's physiological behaviour and performance. The rest of the variables affecting driver's response were kept unvaried to limit their effect. The study acknowledges the delimitations and suggests future work on those variables in section 77.5.

3.6 Summary

Research methods were presented. Based on the analysis, a positivist philosophy with the nomothetic approach is used to design and construct the structural procedure of the study. The study used a driving simulator lab experiment, a questionnaire, and a spiral system development methodology to develop predictive models for the driver's performance measures. Each method was explained in detail and justified, see section 3.2.

The experiment design of the study was presented. Participants were recruited within 20-30 years of age with a minimum of two years of driving experience. The driving scenario was designed and implemented on a STISIM driving simulator. Two NDR tasks were chosen: TQT and email to impose cognitive and visuo-cognitive workload on drivers. The study used a repeated measure approach including three groups, two of them had an NDR task the third had a control group that was not distracted by any NDR tasks. The ethical and experiment procedures were explained and justified.

The data analysis was explained in detail. The study laid out a brief literature review for the data analysis approaches in the literature. Then, a machine learning based approach was suggested and justified. It included feature-based machine learning regressors and classifiers in addition to a multivariate time series classifier. The study used a data preparation technique inspired by previous work to generate a feature subset that provided the highest possible information gain to the ML classifiers and regressors. This included feature generation and selection processes that were designed for the study. Finally, the delimitation of the study was provided.

4

CHAPTER

Using physiological changes to determine the quality of a takeover in Highly Automated Driving

4.1 Introduction

In this Chapter, results of the experiment are explored. The results are concerned with the understanding of how highly automated driving environment affects the physiological data of drivers. The results explore the effect of the NDR tasks on drivers and identifies correlations with drivers' responses. Takeover quality in this study was assessed using two new performance measures called PerSpeed and PerAngle. They are identified as the mean percentage change of vehicle's speed and heading angle starting from a take-over request time using linear mixed models. As outlined in Chapter 3, the physiological measures of drivers were collected. The study focused on heart rate and pupil diameter to explore the effect of highly automated driving on driver's physiological behaviour. In the next sections, driver's physiological changes are explored from the start of the scenario until the end of each takeover. The study also explored the physiological changes imposed by the NDR tasks.

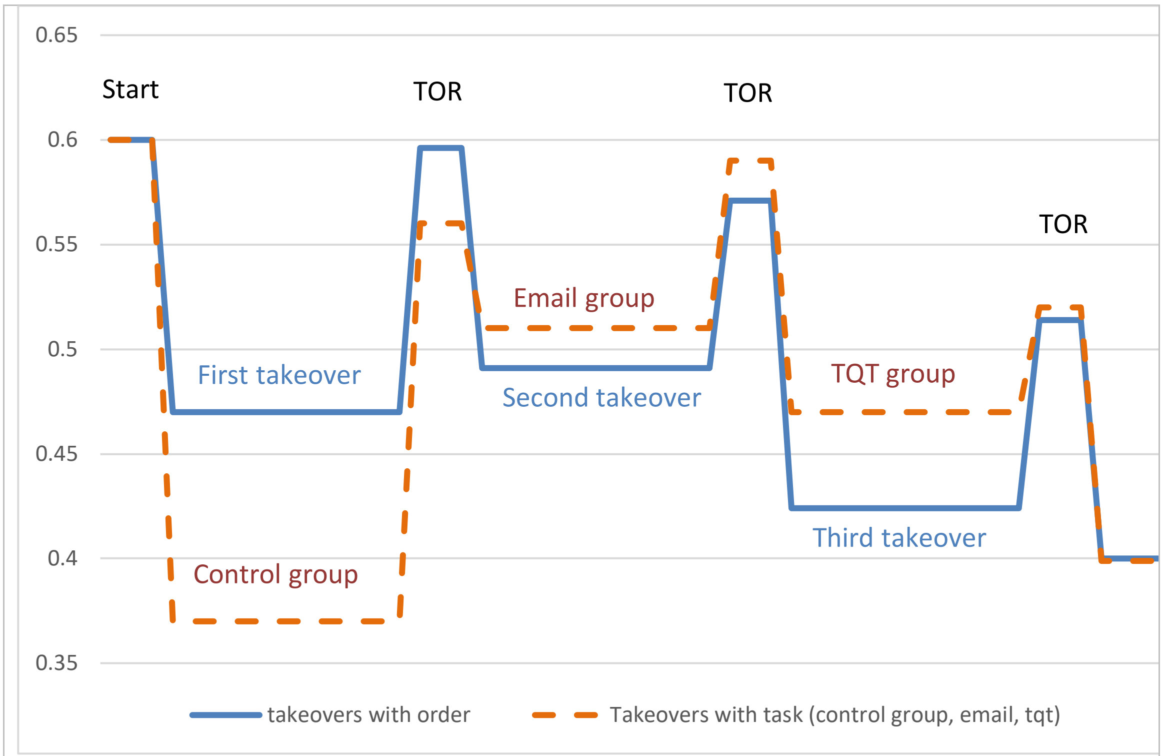


Figure 17: Average normalised HR fluctuations throughout the scenario of the experiment based on 60-second window average. In blue, are the average heart rate values of participants grouped by order of the takeover. In orange (dashes), are the average heart rate values of participants grouped by the secondary task type they performed.

4.2 Physiological Behaviour at the Start of the Scenario

At the beginning of the experiment, a peak in normalised HR and PD were observed in all participants. This was probably due to the stress caused by the anticipation of the experiment. After three seconds of manual driving, it dropped by an average of 20% (M=.6, SD=0.15), see Figure 17, and stabilised (M=.29, SD=.14) as the vehicle switched to automation. To simplify the plot, the aforementioned drop in heart rate after the start is not plotted in Figure 17.

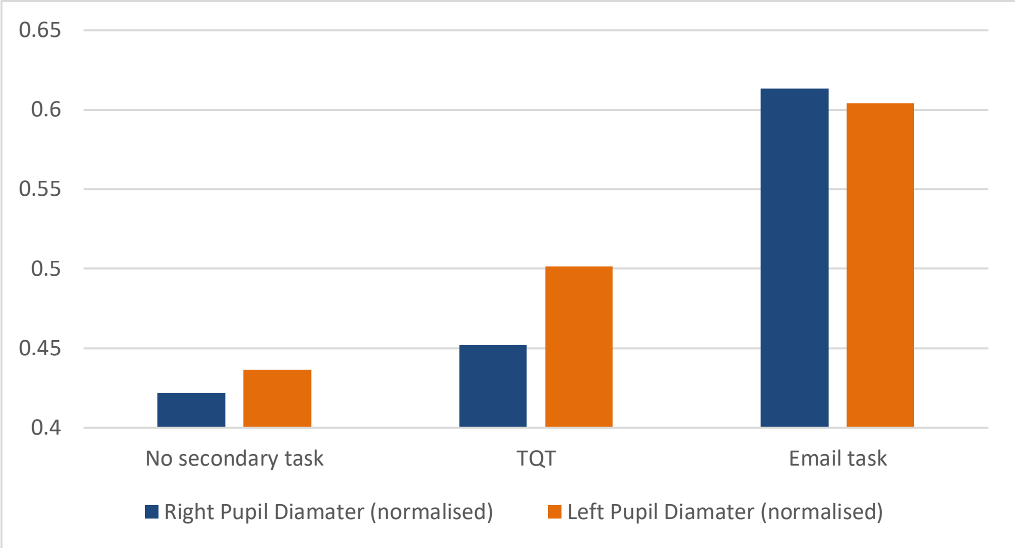


Figure 18: pupil size changes according to secondary tasks.

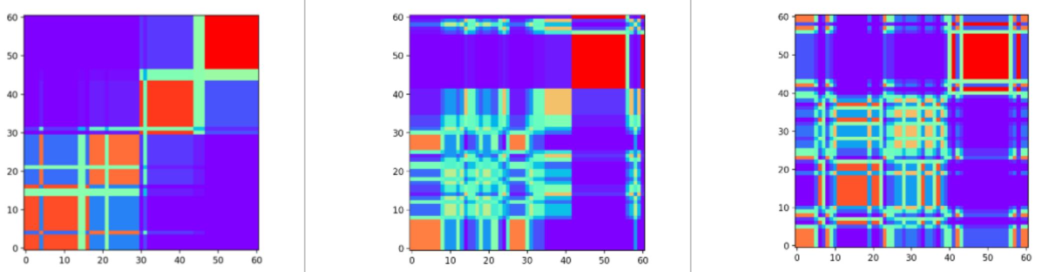
4.3 Effect of NDR Tasks on the Physiological Behaviour of Drivers

4.3.1 Pupil diameter and NDR task

Repeated measures (RM) ANOVA analysis among the three groups of NDR tasks was conducted to explore their impact on the pupil diameter. The data were collected whilst drivers were performing a secondary task prior to a takeover request. There was a statistically significant difference at $F(1.22, 21.9) = 60.741$, $p < .0001$] with Greenhouse-Geisser correction. Post-hoc tests using Bonferroni correction demonstrated an increase in pupil diameter by an average of 0.314 during email task ($p < 0.0001$) and 0.06 during TQT ($p = .02$) in comparison to the control group. A five-second window length demonstrated the highest p values of repeated measures ANOVA test on periods from 1 to 150 seconds. The overall mean values of PD of all participants (after normalisation) are plotted in Figure 18.

4.3.2 Heart Rate and NDR Tasks

The relationship between normalised heart rate and response time was investigated using repeated measures ANOVA to investigate whether there was a significant difference among different groups. Normalised HR, with an average of 90 seconds window, had the strongest significance, $F(2,36)=7.75$, $p<.02$, among window sizes. Pairwise comparison using Bonferroni correction showed a significant difference between the control group and the email task ($p=.03$), the control group and the TQT task ($p=.013$). Differences showed that heart rate increases significantly when drivers are engaged in secondary tasks which align with the results of Carsten et al., (2012). No statistical difference was reported between TQT and email tasks ($p=.95$). Therefore, normalised HR could be considered a valid physiological measure to identify engagement in secondary tasks; however, it cannot distinguish secondary task type for the tasks used in the study reported here.



a. Control Group | b. Email Task Group | c. TQT Group

Figure 19: Markov Transition Field of heart rate of drivers' groups. Colours and squares represent the shapes and transitions of the timeseries. A clear difference is seen between the control group (a) and secondary task groups (b and c). Group B and group C show a slight difference among their signals. Those subjective observations match with repeated measures ANOVA results reported earlier in the study.

4.3.3 The Fluctuations of Heart Rate During the NDR tasks.

Markov Transition Fields (MTF) is one of the recent approaches to encode time series to an image. MTF images "represent the first order Markov transition probability along one dimension and temporal dependency along the other", (Wang and Oates, 2014). MTF images were used to visualise HR signals and their temporal dependencies in an image. So, the more squares and more colours in an image, the more transitions and temporal dependencies are among the signals. It is essential to understand that MTF provides a subjective comparison between signals and is used in this study for this purpose.

To explore the difference in HR transitions among the three groups, MTF images were generated to plot the average transitions (see Figure 19) of the HR signals of each group. Images demonstrated significant signal transitions among TQT and email tasks in comparison to the control group. A clear difference could be identified among the three images. Looking specifically at email and TQT, the signals had similar probabilities of transitions; however, temporal dependencies were significantly different. In contrast, the control group had fewer transitions probabilities and less temporal dependencies. This meant that during the NDR tasks, HR transitions were much more frequent than the control group. In addition, the collected HR fluctuations during the TQT were much higher than the email group.

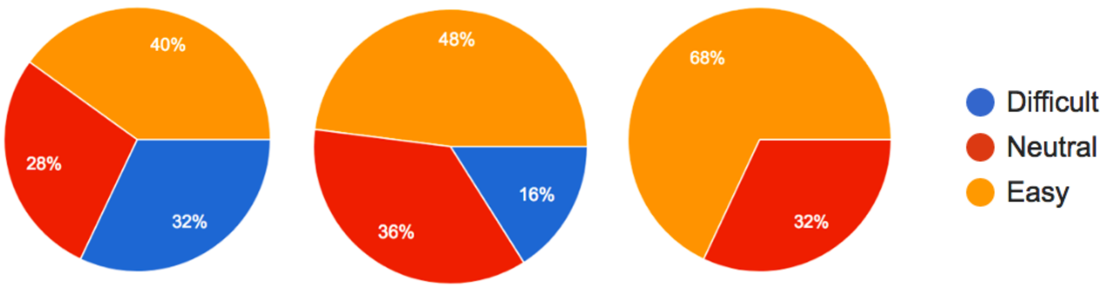
4.4 Effect of the Transition on the Physiological Behaviour of Drivers

Takeover requests sparked a peak in HR (M=.43, SD=.2) that lasted for few seconds, see Figure 17. Then, HR gradually drops to the mean of the control group (M=.36, SD=.14) among the three groups within a mean of 10 seconds (SD=6.2) of a successful takeover. In takeovers that ended with an accident, the significant increase of HR remained for a more extended period of time (M=20, SD=6.2).

4.5 Effect of the Order of the NDR Tasks

Design of repeated measure studies could cause a severe ambiguity and bias when learning or practice effects are not taken into account; especially when observations of learning effects are reported in the handover process studies (Larsson et al., 2014). The repeated measure-approach was used in this study to assess whether the reported learning curve may affect the physiological behaviour of drivers; specifically, their heart rate changes.

Consequently, one-way analysis between takeover groups based on the analysis of variance was performed to assess the correlation between heart rate and the order of takeover requests. Average heart rate was calculated of all participants grouped by the order in which a takeover is performed. For example, the second takeover group means the values collected at the second takeover among all participants regardless of the secondary task type they performed.



a. First handover b. Second handover c. Third handover

Figure 20: Subjective ratings of the difficulty of first, second and third handovers.

Results showed that the second takeover group had a higher HR mean (M=.49, SD=.13) than the first (M=.462, SD=0.13) and third (M=0.42, SD=0.14) groups. There was a difference (F=3.1, p<.05) even though the difference in mean between the three groups was quite small. Therefore, the hypothesis that HR could decrease over time because of drivers experience in handling the takeover was rejected. Accordingly, the study suggests that the order of the secondary tasks has no significant influence on the physiological changes happening during the secondary task engagement.

4.6 Subjective Ratings of Takeovers' Difficulties

At the end of the experiment, participants reported the difficulties of each takeovers, as reported in Figure 20. Participants reported the first takeover as the most difficult one among the second two takeovers. It's apparent that participants perception of difficulty decreases gradually. This could be interpreted as a confirmation of the existence of the learning effect on the driver's perception and performance on the takeover performance.

4.7 Response Time Estimation Using HR and PD as Covariates

Due to the repeated measure approach of the study, the statistical model used to assess the correlation between response time, HR and PD was a repeated measure linear mixed model (LMM). Since LMMs assume their covariate variables to be independent, left pupil diameter was excluded from this analysis due to its strong dependence with right pupil diameter, $r=.98$, $p=.002$.

On the main effects (at window size=30s), right pupil diameter was significant, $F(1, 30.5)=12.2$, $p=.001$, and HR was not significant $F(1, 27.2)=3.8$, $p=.06$. When using secondary task type as a fixed effect in the model, HR demonstrated a strong significance, $F(1, 30.3)=11$, $p=.002$. This means HR window length may not have been long enough to provide enough significance among groups. Moreover, HR and right PD had a strong interaction term, $F(1,24.7)=4.2$, $p=.049$. Table 10 shows the interaction effect among variables and intercepts value of the analysis. This confirms that normalised HR has a predictability potential for response time assuming the right window size is chosen.

Table 10: Estimates of Fixed Effects on response time using a 30-sec window size

Parameter	Estimate	Std. Error	Sig.
Intercept	1.100432	.361619	.005
hr	1.077343	.552368	.062
pdr	6.786712	1.935462	.001
hr × pdr	-4.645041	2.247214	.049

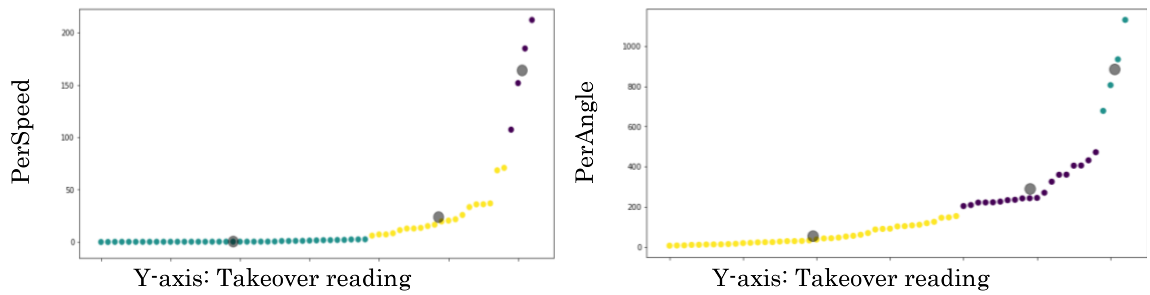
In order to understand if window size influences the analysis, the same statistical methods were applied to the data extracted from a 60 second window size. HR was statistically significant $F(1,24.3)=6.2$, $p=.019$ with no secondary tasks type added as a fixed effect. When added, significance increased, $p<.0001$. Conversely, right pupil diameter was not significant, $F(1, 19.2)=1$, $p=.317$. This indicated that window size has a strong effect on the correlation between physiological changes and response time since pupil diameter was significant at 5-second window size.

In order to find the optimal window size in which heart rate and pupil diameters performed at, a simple optimisation algorithm was run. Results, as indicated above, showed that a 30 seconds window was the best performing window size for PD correlation with response time and 60 seconds for heart rate. This could be correlated to HR responding slower than PD to external changes. Hence, a longer HR window captures long-term physiological changes, and PD captures short-term physiological changes; hence window size values reacted accordingly. Those findings align with Solovey et al., (2014) that reported that the best window size for their physiological data was 30 seconds.

4.8 Quality of the Takeover

Response Time analysis examines the readiness to respond but not necessarily the quality of drivers' responses. Response time is not the only measure of performance in takeover situations; other studies take the quality of takeover as another important performance measure (Gold et al., 2016; Radlmayr et al., 2014; Zeeb et al., 2016).

Few studies used minimum time to collision as a performance measure (Gold et al., 2016; Körber et al., 2016; Radlmayr et al., 2014); however, sharp changes in speed or heading angle of the vehicle are considered poor performance indicators in motorway driving scenarios (Kass et al., 2007; Zeeb et al., 2016); especially in this study's scenario design. To evaluate that, two more performance measures were constructed from the collected data; those variables were used to assess the correlation between driver's HR and PD changes (pre-TOR), and the quality of the response was investigated accordingly.



a. b.

Figure 21 :Mean Percentage Change of a.) Speed and b.) Heading angle of the vehicle. Participants are split independently into three clusters representing low, medium and high-risk groups. Black circles represent the centroid of each cluster.

4.8.1 Understanding PerSpeed and PerAngle Variables

To quantify the changes in vehicle speed, *PerSpeed and PerAngle* measures were introduced in this study. PerSpeed is the mean percentage change of vehicle's speed for a period before the TOR (e.g., 30 seconds), see Equation *3*. A higher percentage indicates a sharper change in speed which could be either braking or acceleration. *PerAngle*, similarly, is the mean percentage change of vehicle's heading angle. In this study, no rapid change in speed or heading angle was required because the driving scenario allowed participants to perform a smooth transition to the next lane. Therefore, suczh actions were penalised in this study because it correlates with a lack of situational awareness according to Kass et al., (2007). PerSpeed and PerAngle at each takeover are plotted in Figure 21 by sorting takeover incidents by their corresponding PerSpeed (a) and PerAngle (b) values.

Table 11: Definition of clusters of PerSpeed and PerAngle.

	PerSpeed				PerAngle			
	Start Range (%)	End Range (%)	HR mean	HR std.	Start Range (%)	End Range (%)	PDR mean	PDR std.
Group 1	0	20	0.47	0.13	0	195	0.5	0.2
Group 2	>20	40	0.56	0.10	>195	453	0.7	0.08
Group 3	>40	220	0.61	0.11	>453	1200	0.57	0.19

The K-Means clustering algorithm (MacQueen, 1967) was used to cluster PerSpeed and PerAngle readings independently into three groups. Since Zeeb et al.'s, (2016) study split participants into three groups based on the quality of their response, this study followed the same approach by defining clusters 1, 2 and 3 as low, medium and high-risk groups. Clusters start/end ranges identified by the K-Means algorithms, and their corresponding mean HR and PD are defined in Table 11. The table includes start and end ranges of both HR and PD of each cluster, their mean and standard deviation values. As indicated in Figure 21, the high-risk group had four incidents; three of them ended with an accident, and the fourth one was an anomaly where a participant decided to stop the vehicle for six seconds before deciding to move and change lane.

$$y = \frac{100}{n} \times \sum_{i=k}^{n} \frac{x_{\text{i}} - x_{i-1}}{|x_{i-1}|}$$

Equation 3: Mean percentage change formula, where y is PerSpeed or PerAngle, x is a vector containing readings (speed or heading angle values), k is time$_0$, and n is window size.

Statistical analysis of K-Means groups based on their corresponding heart and pupil diameter data yielded interesting results. A one-way ANOVA showed that the mean normalised HR had a significant difference, F=4.2, p=.01, among PerSpeed groups. These findings indicate that higher HR means a higher probability of strong braking; which is considered a bad performance measure. No significant difference was reported based on drivers' PD mean values.

When analysing PerAngle K-Means groups, no correlation, F=2.4, p=0.09 was identified between PD and the mean of low (m=.5, SD=.2), medium (m=.7, SD=.08) and high (m=.57, SD=.19) risk groups. Though, higher risk groups had higher pupil dilation than the low risk group. Based on Batmaz and Ozturk's, (2008) findings, pupil diameter dilates with the mental workload which explains the increase of PD mean value from low to high-risk groups. However, the medium risk group had a significantly higher mean than the high risk one.

The deviation of the medium risk group may be explainable because the main NDR task performed by the medium risk group was the email task. It was performed by 63% of participants of that group. Reflecting that on the findings, such an increase in PD in the second task group could have been due to the difference in lighting between the tablet screen and the simulator screens. This could explain the significant difference in pupil diameter in comparison to other groups; even though, the experiment setup ensured a minimal change in lighting throughout all screens by manually setting all brightness on the screens. Another explanation could be due to the change in pupil diameter as participants transit from one screen to another since the lab was significantly darker than the two screens.

The results align with the established findings in the literature that the pupil diameter measure is valid only in highly controllable environment (Marquart et al., 2015). This means that pupil diameter estimation in real-world driving may not be accurate; however, there has been significant research in assessing mental workload under different lighting conditions (Pfleging et al., 2016). Hence, results presented in this study are potentially useful in real-world applications.

4.8.2 PerSpeed and PerAngle Analysis

To measure whether the study's task type have any influence on PerAngle and PerSpeed, a linear mixed model using Toeplitz covariance type with repeated measures test was performed. The task was used as a fixed effect; HR and PD were used as a covariate to understand whether they significantly impacted the predefined quality measures. Tests were done using linear mixed models.

For PerSpeed performance metric, task type was significant $F(2, 43)=4.3$, $p=.019$, as was HR, $F(1,44)=5.5$, $p=.01$, and PD, $F(1,46)=2.5$, $p=.01$. All other higher-order interactions were significant, specifically PD × HR, $F(1,46)=8.2$, $p=.006$ that had the highest significance. The Bonferroni test showed no significant differences in PerSpeed values among task groups. Additionally, estimates of fixed effects demonstrated that each one per cent increase in HR corresponds to 4.6% decrease in PerSpeed ($p=.002$) and for each one per cent increase in PD corresponds to 9.1% decrease in PerSpeed ($p=.004$).

Similarly, PerAngle metric, the secondary task type, $F(2,18)$, $p=.001$, was significant, so was PD, $F(1,26)=4.4$, $p=.04$, and HR, $F(1,26)=5.1$, $p=.003$. Higher level interactions were not more significant, and Bonferroni pairwise comparison showed no difference among task groups. Estimates of fixed effects indicated that one per cent change in HR corresponds to 54% change in PerAngle and for each one per cent in PD, 71% change in expected. Such results indicate that physiological measures are valid predictors for PerSpeed and PerAngle performance measure.

The results demonstrate further evidence that HR and PD correlate with braking behaviour of drivers. According to (Gold et al., 2013a), braking is associated with out-of-loop drivers allowing themselves a longer time to restore situational awareness. Due to an increase in mental workload is associated with an increase in HR (Wilson, 2002) and PD (Batmaz and Ozturk, 2008), the reported results indicated that PerSpeed, PerAngle and driver's mental workload have an indirect negative correlation; assuming drivers are out-of-loop.

4.9 Summary

The most significant findings are, heart rate and pupil diameters of drivers are valid predictors for both response time and determining the quality of takeovers in highly automated driving environments. Interestingly, these results are similar to findings in air traffic control and aviation systems in addition to manual driving studies that were performed previously. The findings of this experiment paved the path to assess the possibility to predict the response time and the quality of takeover prediction models. The models could be applied to all drivers based on the physiological behaviours without necessarily accounting for individual differences or relying on identifying the secondary tasks drivers were performing.

Moreover, two new quality measures were introduced and examined in this study to provide an estimate of braking and steering, and they were linked to drivers' physiological measures. PerSpeed and PerAngle measures could be used by the automated driving systems to assess the driver's future responses. In addition, PerSpeed and PerAngle showed a strong correlation with driver's NDR tasks when used as fixed factors in mixed linear models. The study split drivers into low, medium and high-risk groups by applying K-Means clustering algorithm on PerSpeed and PerAngle values. The heart rate of drivers had a strong correlation with driver's clusters of the PerSpeed value. Overall, the study established that heart rate and pupil diameter values have a strong correlation with driver's braking style. The reported results provide insight into driver's readiness and therefore, allow automated systems to adopt the right driving strategy and plan to enhance their experience and make the transition phase safer for everyone.

5

CHAPTER

Predicting Drivers' Performance in Highly Automated Driving

5.1 Introduction

In highly automated vehicles, drivers can deviate their attention from the road and engage in non-driving related tasks whilst the car is driving itself. When the automated system prompts drivers to take over the driving task to handle a critical hazard, drivers have a lag time when they are trying to restore situational awareness before acting (Merat et al., 2012). That is called response time and has been found to be influenced by tasks being performed before the takeover request. Previous studies showed that response time is not always correlated with response quality, i.e., driver's ability to manoeuvre safely from a critical hazard.

The aim of this study was to assess the predictability of response time and quality based on drivers' physiological measures before the takeover request. The dataset was collected during a driving simulator study presented in Chapter 3. The time-series of heart rate, pupil diameter and saccade eye movements were used to generate sets of features. Then, those features were the input of statistical and machine learning algorithms to predict response time and quality of drivers' takeovers. Machine learning results using classification and regression methods are presented.

The machine learning models aimed at classifying drivers' responses. A time-series based classification was also explained and evaluated. Results of the classification of response time and quality are outlined in the next sections.

5.2 Results Using Regression Methods

The study aimed at predicting response time and the driving performance quality measures: PerSpeed and PerAngle. The regression methods used were linear regression algorithm, random forest regressor and bagging regressors, see 3.4.7.2 for more detail. Results for each performance methods are detailed in the next subsections. The regression had three target variables that were assessed independently, response time, PerSpeed and PerAngle. Results of each one of them are reported in their corresponding sections.

5.2.1 Response Time

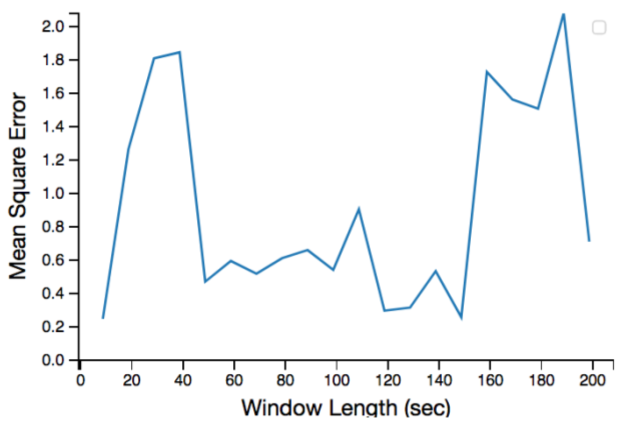


Figure 22 Mean square error of the linear regression method (Lasso).

First, a linear regression algorithm was used with L1 prior as regularizer (known as Lasso) (Tibshiranit, 1996) to find a linear correlation between extracted features and the response time (in categories). Extracted features are either simple or relevant, as explained in section Features Selection. Results show, see Figure 22, sufficient evidence of the accurate predictability of the model at a window size of 150 seconds. Analysing the data shows, windows of 120- and 150-seconds lookback time had a mean squared error (MSE) of 0.23 – this value reflects a good accuracy of the prediction of the response time. However, a higher MSE error was observed at other window sizes. This variability in MSE caused by the window size was observed in similar studies (Grimes et al., 2008; Solovey et al., 2014). Since linear regression models don't explicitly specify why specific window sizes perform better, some effort was spent in identifying similar results in other studies to provide more insight into this.

Table 12: Prediction performance of regression models.

Model Name	Output	Features	Window Size (Best)	Results (MSE)	
				Train	Test
Linear Regression using Lasso	Response Time	Simple	120	1.9	1.8
		Relevant	150	0.21	0.23
	PerSpeed	Simple	60	35.0	59.1
		Relevant	60	32.1	35.0
	PerAngle	Simple	90	21.0	25.0
		Relevant	60	19.3	24.9
Random Forest Regressor	Response Time	Simple	150	1.3	1.6
		Relevant	90	0.23	0.24
	PerSpeed	Simple	40	36.0	54.1
		Relevant	90	30.1	49.0
	PerAngle	Simple	90	19.5	24.0
		Relevant	120	19.3	24.9
Bagging regressor with random patches (based on decision tree estimator)	Response Time	Simple	90	2.4	2.7
		Relevant	150	0.18	0.19
	PerSpeed	Simple	120	25.0	29.1
		Relevant	60	24.1	24.8
	PerAngle	Simple	90	25.0	38.0
		Relevant	120	17.9	19.2

The training data has a set of features over a specific time window; each window has a significant feature that could have an impact on the response time. For example, few studies have identified mean pupil dilatation of a window size of 30 seconds to

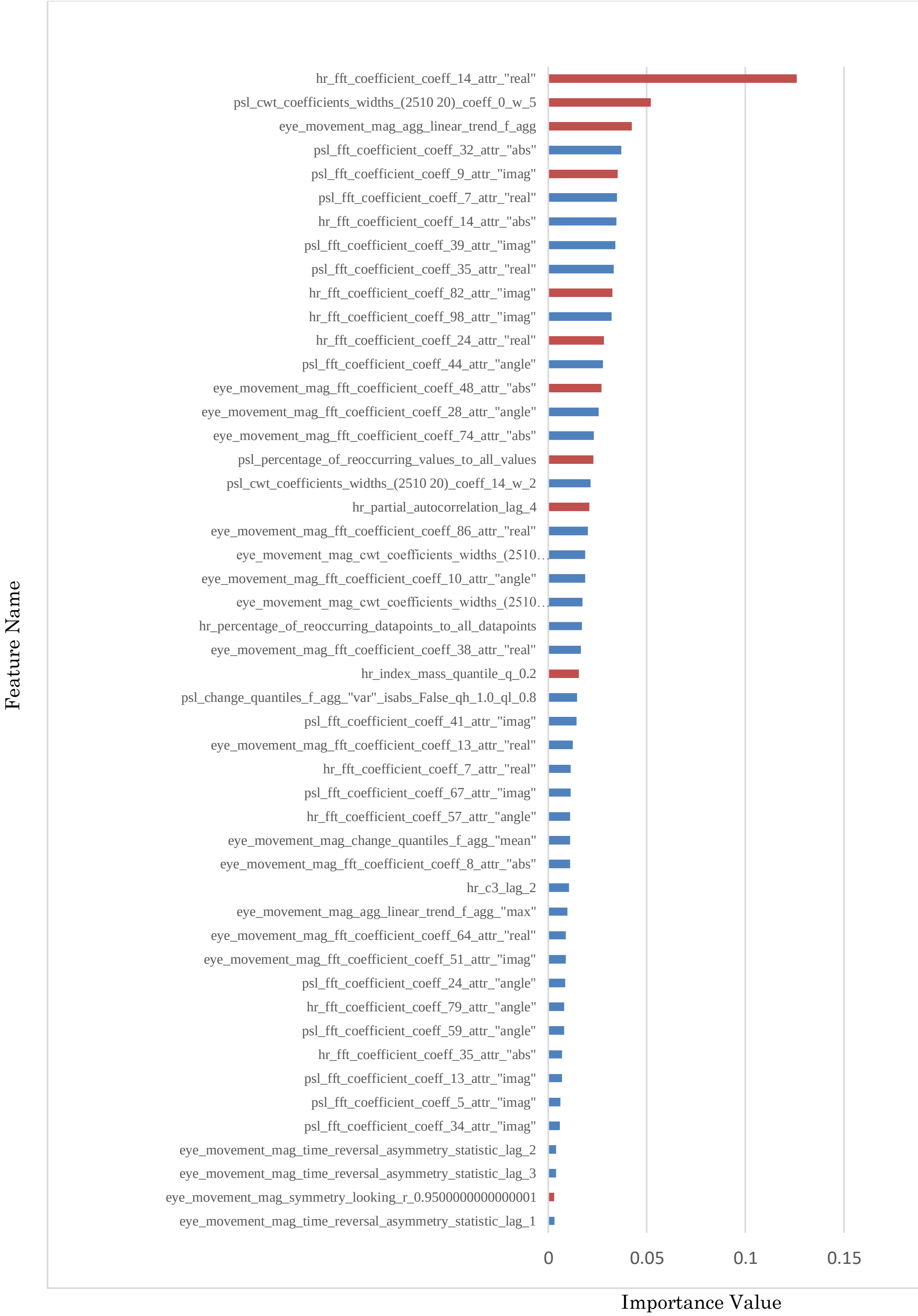


Figure 23: List of features passed to the response time regression model. Features are ranked by the random forest information gain values. Features highlighted in orange formed a subset that had the highest information gain with the least collinearity; making the best combination of features for an ML predictor.

have the highest statistical significance to mental workload (e.g. Grimes *et al.*, 2008) among other windows sizes.

Based on the regression results, physiological changes happening between 20 and 120 seconds before a TOR have a linear correlation with response time. However, changes occurring before 20 and after 120 seconds may not have a linear correlation. Figure 23 shows a list of features with the highest information gain; those features were the input of the regression model. Most of those features are on the frequency domain which means their values are highly dependent on the window size. Therefore, the regression model is expected to behave differently. Features and their generation process were discussed in section 3.4.8.5.

Current literature has limited research on the effect of heart rate on response time in highly automated driving . So, seeking a better understanding of this behaviour through literature is not possible. The results show that the windows smaller than 20 seconds don't have sufficient features to feed the regression model which is evident by the results of the experiment. On the contrary, windows over 150 seconds imposed a noticeable lag over the supplied information to the model. To explain this further, events happening beyond 150 seconds may have no influence on the driver's mental state at the takeover time; which is 150 seconds later. As a result, any physiological data beyond that time biases the model. Since the dataset is quite limited, the explanations provided may not be entirely accurate and more research is required in this area.

The second approach used was ensemble models which showed few interesting results. First, ensemble models had no significant impact on the regression accuracy, see Table 12 for both simple and relevant features. Random Forest ensemble approach showed better MSE on train set (1.3) but less accuracy on the test set (1.6) using the simple features set. Still, that approach was less accurate than the simple regression approach. However, Random Forest regressor performed nearly the same as Lasso linear regressors with only 0.1 difference in MSE.

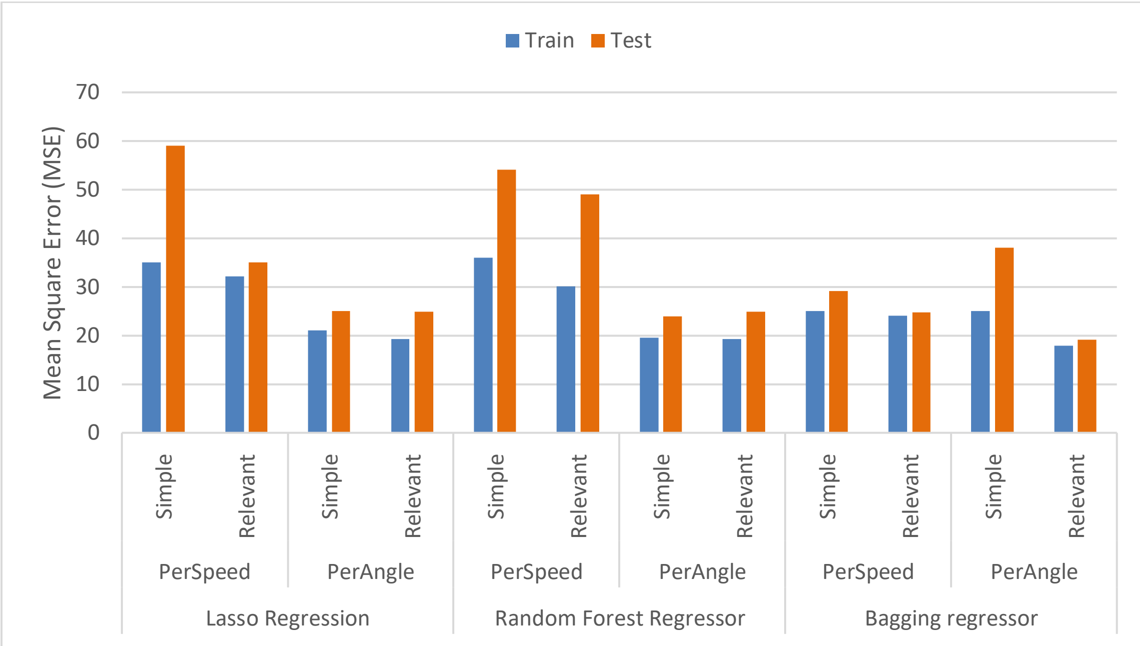


Figure 24: An illustration of the difference in MSE values in train and test datasets. Results show a strong deviation between the train and test MSE. This deviation indicates that the ML models were overfitted to the training data.

Bagging regressors with random patches using decision trees outperformed the aforementioned approaches. Again, relevant features had an MSE of 0.19 on the test set at window length of 150 seconds. It's interesting to mention here that relevant features were consistently better than simple features at generating better regression models using the two approaches. Though, generating all features and filtering them was computationally more expensive than simple features generation which in turn provides slower and computationally demanding algorithms. To explain, relevant features processing took approximately twenty minutes on a MacBook Pro Core i7 using four threads to compute them in parallel. In comparison, simple features took 300 milliseconds on the same machine.

5.2.2 Performance Quality Measures

PerSpeed and PerAngle had a high MSE in all regression models. Differences in MSE values between train and test accuracies showed a strong tendency to overfit the training data, see Figure 24. Most of the generated regression models converged by outputting the mean value of the entire output for PerSpeed or PerAngle regardless of the input values. This could be due to the high dimensionality of the data or a non-linear relationship between the features and the performance measures. It's also noticeable that bagging regressor had a minimal deviation between training and test datasets; however, the MSE values were between 18 and 22 which is a high error. Due to the poor results of the regression methods, classification-based methods were tested.

5.3 Results Using the Classification Methods

A classification approach was used to assess the predictability of the performance quality measures, e.g., response time, PerSpeed and PerAngle. In this section, seven different classification algorithms are presented, and their results are discussed. Implementation of those algorithms was provided by Scikit-Learn, a machine learning Python library (Pedregosa et al., 2011)

Reported classification sufficiency is reported in f1-score, precision and recall values. Precision is defined as the ratio of relevant instances among the retrieved instances. The recall is defined as the ratio of relevant instances that have been retrieved over the total amount of relevant instances. F1-score is the harmonic average of precision and recall (Davis and Goadrich, 2006).

Table 13: Performance measures of classification methods.

Model Name	Output	Features	Window Size - Best	Results - Train		Results - Test		
				Precision	Recall	Precision	Recall	F1
Random Forest	Response Time	Simple	30	0.93	0.89	0.89	0.87	0.88
		Relevant	40	0.91	0.95	0.92	0.94	0.93
	PerSpeed	Simple	60	0.95	0.9	0.92	0.89	0.90
		Relevant	60	0.81	1.0	0.94	0.98	0.96
	PerAngle	Simple	90	0.87	0.8	0.81	0.82	0.81
		Relevant	60	0.8	0.82	0.86	0.85	0.85
1-Nearest Neighb-our	Response Time	Simple	30	0.89	0.72	0.75	0.7	0.72
		Relevant	40	0.93	0.83	0.82	0.84	0.83
	PerSpeed	Simple	40	0.93	0.92	0.89	0.87	0.88
		Relevant	90	0.8	0.95	0.92	0.94	0.93
	PerAngle	Simple	90	0.81	0.86	0.81	0.85	0.83
		Relevant	90	0.43	0.9	0.82	0.88	0.85
Gradient Boosting	Response Time	Simple	40	0.44	0.41	0.71	0.72	0.71
		Relevant	40	0.83	0.48	0.73	0.75	0.74
	PerSpeed	Simple	120	0.86	0.8	0.82	0.82	0.82
		Relevant	120	0.81	0.84	0.86	0.83	0.84
	PerAngle	Simple	90	0.87	0.8	0.71	0.72	0.71
		Relevant	120	0.83	0.82	0.76	0.75	0.75
Decision Tree	Response Time	Simple	30	0.91	0.81	0.81	0.8	0.80
		Relevant	30	0.91	0.91	0.9	0.95	0.92
	PerSpeed	Simple	30	0.92	0.89	0.91	0.89	0.90
		Relevant	30	0.78	0.95	0.93	0.95	0.94
	PerAngle	Simple	60	0.83	0.8	0.71	0.81	0.76
		Relevant	90	0.8	0.84	0.83	0.8	0.81
SVM	Response Time	Simple	60	0.89	0.85	0.91	0.9	0.90
		Relevant	60	0.84	0.85	0.95	0.92	0.93
	PerSpeed	Simple	30	0.89	0.83	0.83	0.83	0.83
		Relevant	120	0.8	0.83	0.82	0.83	0.82
	PerAngle	Simple	90	0.88	0.84	0.83	0.81	0.82
		Relevant	120	0.40	0.81	0.89	0.82	0.85
Naïve Bayes	Response Time	Simple	120	0.40	0.41	0.43	0.41	0.42
		Relevant	100	0.83	0.48	0.41	0.4	0.40
	PerSpeed	Simple	120	0.81	0.8	0.82	0.82	0.82
		Relevant	120	0.81	0.82	0.82	0.81	0.81
	PerAngle	Simple	120	0.84	0.8	0.81	0.82	0.81
		Relevant	120	0.93	0.82	0.83	0.81	0.82
MUSE + WEASEL	Response Time	Timeseries	N/A	0.91	0.83	0.56	0.5	0.53
	PerSpeed	Timeseries		0.95	0.83	0.4	0.42	0.41
	PerAngle	Timeseries		0.4	0.94	0.51	0.3	0.38
Random algorithm	Response Time	N/A	N/A	0.4	0.3	0.4	0.3	0.34
	PerSpeed			0.2	0.4	0.5	0.5	0.50
	PerAngle			0.21	0.39	0.4	0.45	0.42

5.3.1 Response Time

First, a decision tree with Gini classifier was used. Decision Trees perform a classification based on the frequency of labelled variables to maximise the information gain (Raileanu and Stoffel, 2004). Decision tree results have shown that f1-score was an average of 0.81; nearly the same value was observed across all window sizes.

Second, a Random Forest classifier was used. It's an ensemble approach to construct classifiers by merging multiple decision trees at training time and outputting the most common output among all trees (Breiman, 2001). To avoid overfitting, 40% of the data was used as a test set. Overall f1-score of Random Forest classifier was 0.88 for simple features set and 0.92 for relevant features set. It's noted that a window size of 40 seconds had the highest f1-score.

Third, Gradient Boosting Classifier is a boosting machine learning classifier that uses an ensemble of decision trees that incorporates a loss function for better optimisation than random forest (Friedman, 2001). It iteratively builds a large set of shallow decision trees to reduce bias (Freidman, 2008). Each decision tree is trained to improve the output of the previous one (Friedman, 2001). The overall f1-score of the algorithm was approximately 0.71 which is significantly lower than random forest results. This difference could be due to the small size of the data set even though a previous study reported that Gradient Boosting algorithms perform efficiently with small datasets (Zhao et al., 2019).

Fourth, the 1-Nearest Neighbour classifier was used to assess if there's a pattern among the classes. It's an instance-based learning method that finds the closest k neighbours to the instance in the training data (Cover and Hart, 1967). Even though it's among the most straightforward machine learning algorithms, it outperforms several algorithms in several practical problems (Caruana et al., 2005). Results showed that relevant features set increased the overall f1-score by 11% in comparison to the simple feature set, see Table 13: Performance measures of classification methods. A standard deviation of 0.13 is observed among window sizes where it performs best at window sizes 140 and 170 with f1-score of approximately 0.82.

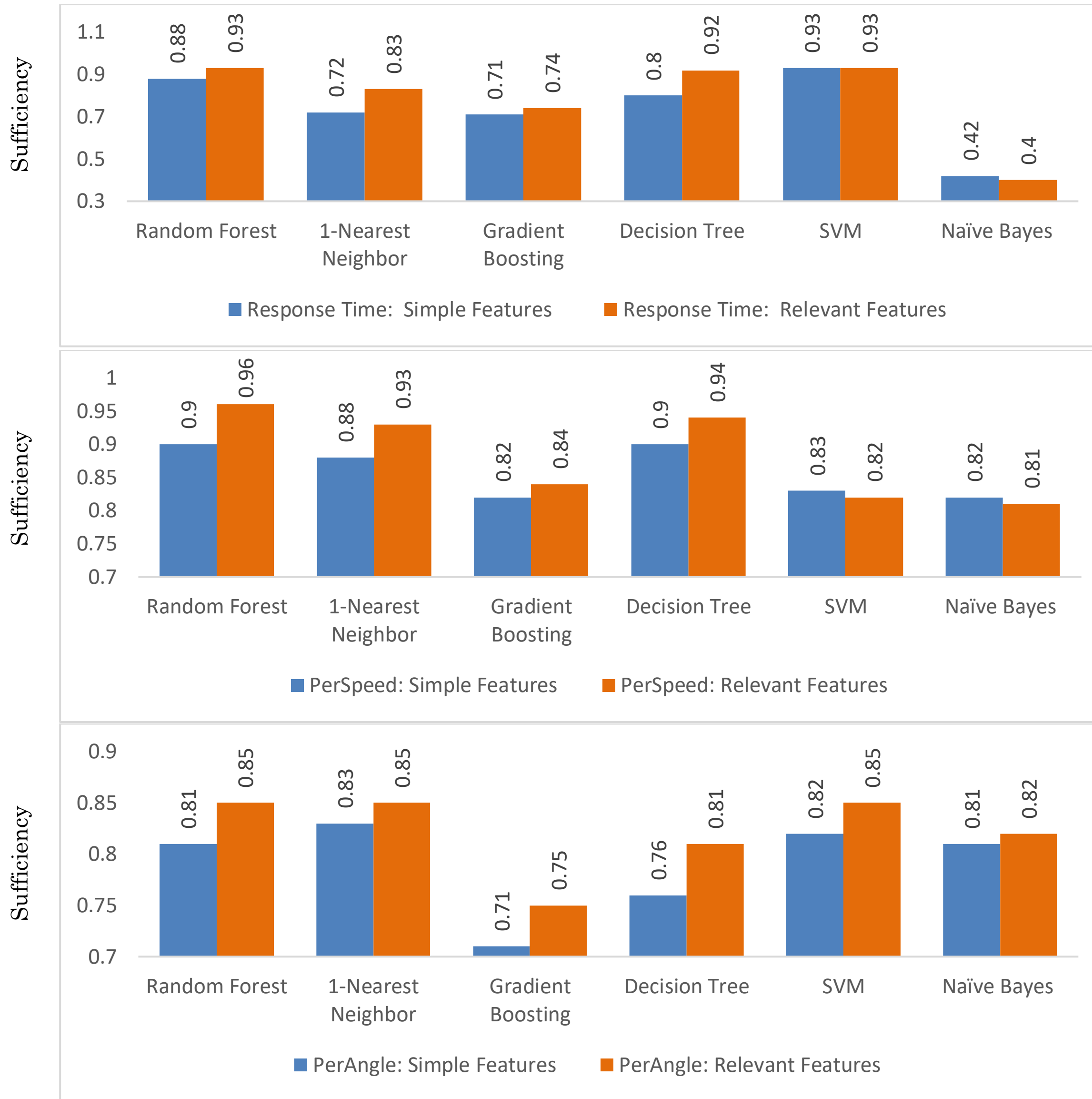


Figure 25: F1 score difference between classification algorithms using simple and relevant features sets.

SVM with C-Support Vector Classification with radial basis function (RBF) kernel was used. It's a supervised learning approach that uses the RBF kernel to split up non-linear training data by mapping them to high dimensional feature space

(Suykens and Vandewalle, 1999). Using a hyperplane, it splits up data to n-dimensional classification space to provide the highest margin between classes; this allows better classification and less overfitting than other algorithms (Smola and Schölkopf, 2004). Looking at the results in Table 13, SVM outperformed other reported classifiers in both simple and relevant features with 0.9 and 0.92 for simple and relevant features. This outperformance could be due to SVM's ability to perform with high dimensionality data (Weston et al., 2000) which is evident in this dataset based on the results of other classifiers.

Finally, Table 13 shows the precision and recall values of both train and test results. Most classifiers showed no significant differences between their train and test accuracies. Moreover, all reported algorithms performed significantly better than the random algorithm where f1-score was only 0.35 in comparison to SVM's f1-score at 0.93. It was essential to make such a comparison since the output labels were unbalanced. Results, in Table 13, show that most algorithms didn't get biased by the dominant output classes.

In contrast, WEASEL+MUSE algorithms performed poorly where their f1-score was 0.53 on test data which is significantly lower than all other algorithms, see Table 13. Looking at training data algorithms, a much higher f1-score (0.85) was observed signifying clear overfitting over the training data. WEASEL+MUSE's main drawback was overfitting in addition to the heavy computations required for training as reported by the authors of the algorithm (Schäfer and Leser, 2017a). They also mentioned that the overfitting problem might be mitigated by tuning the hyperparameters of the algorithm. Due to the expensive computations required for each iteration of training WEASEL+MUSE, the hyperparameter optimisation was not considered since other classifiers performed well in response time classification.

5.3.2 Measures of Performance Quality

As mentioned in section 3.4.8.7, PerSpeed and PerAngle values were discretised using K-Means cluster into three groups. Those clusters were used as labels to build classification algorithms. When predicting PerSpeed, 1-Nearest Neighbour, Decision Trees, and Random Forest classifiers had similar performance with f1-score of 0.93, 0.94 and 0.96 consecutively using relevant features set. An average of two per cent increase in f1-score was observed in comparison to simple features set in classifiers. Gradient Boosting, SVM and Naïve Bayes algorithms reported an average f1-score of 0.82. However, MUSE+ WEASEL reported 0.86 for the training set and 0.53 for test set which is an indication that it overfitted the training data. Similarly, PerAngle's best-performing classifiers were Decision Trees, 1-Nearest Neighbour, SVM with f1-score of 0.82, 0.85, 0.85 consecutively.

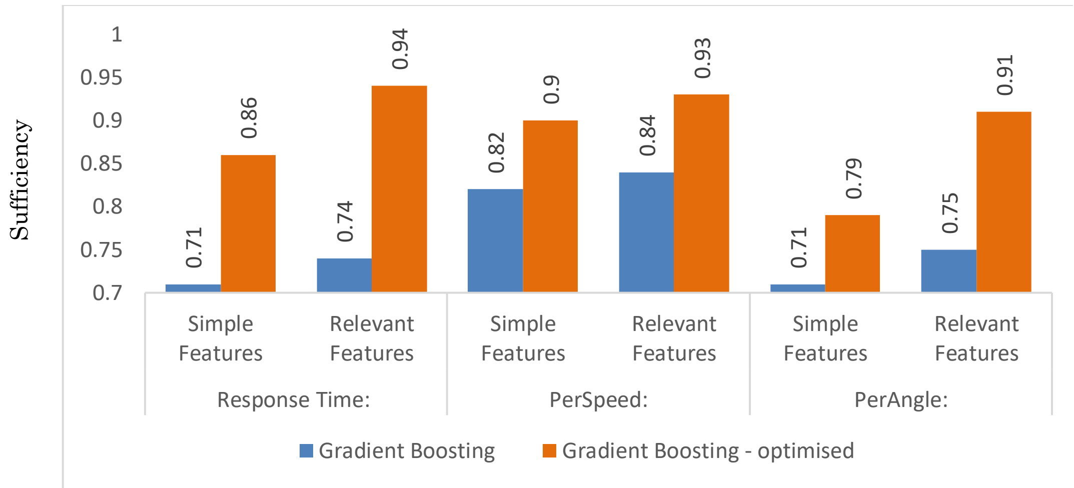


Figure 26: the difference between optimised and default parameters using Gradient Boosting classifiers

Figure 25 shows the f1-score of classification algorithms using both simple and relevant features set. In PerSpeed and PerAngle results, relevant features increased the accuracy of all algorithms with an average of 3.6%. For example, Decision Trees performance increased by 12% in classifying Response Time, 4% in classifying PerSpeed and 5% in classifying PerAngle. Such an increase could be due to the higher information gain in the relevant feature sets. As such, decision trees could reach higher predictability using the same depth of nodes. Interestingly, no significant difference is observed in SVM using relevant or simple features; this could be due to SVM's resilience to collinearity and its ability to separate features on hyperplanes (Bennett and Campbell, 2000).

Gradient Boosting classifier performed poorly in comparison to other ensemble classifiers such as the random forest algorithm. This may have been caused by the high sensitivity of Gradient boosting algorithm to its configuration parameters (Bühlmann and Hothorn, 2007). To assess this, grid search analysis was performed to find the optimal values of tree depth, the number of trees and shrinkage parameters. Results are illustrated in Figure 26. They have shown a significant increase in Gradient Boosting algorithm's performance. Such improvement in accuracy proves that Gradient Boosting algorithms require parameter tuning.

5.4 Conclusion

In this study, regression and classification methods were used to assess the predictability of response time and performance quality measures of drivers in highly automated driving based on their physiological behaviour. Performance quality measures were PerSpeed and PerAngle that were identified in a previous study. Physiological behaviour was represented by three time-series; heart rate, pupil diameter and magnitude of eye movements. A regression method using Lasso achieved accurate predictability of response time with 0.23 MSE. Though, regression approaches had a significantly higher MSE for PerSpeed and PerAngle values.

Therefore, classification-based ML algorithms were used. This study used an evolutionary based approach to find a near-optimal subset of features to feed to machine learning models. Results showed that there was an average of 9.3% increase in f1-score in all trained models in comparison to simple features. This was due to the enrichment of information due to the extensive number of generated features. To evaluate the performances of the selected classification algorithms, the study used precision, recall and f1-scores. 1-Nearest Neighbour, Decision Trees, and Random Forest demonstrated the highest f1-score among all algorithms achieving an f1-score of 0.96 for PerSpeed and 0.85 for PerAngle.

The most important conclusion of this study was that physiological behaviour of drivers before a takeover is sufficient to predict both response time and quality after a takeover request in highly automated driving. The results also showed that the window size of the time series played a significant role in the stability of the algorithms. The results provided few justifications for this behaviour but couldn't back it up with experimentation due to the limited size of the dataset.

6 CHAPTER

Discussion

6.1 Introduction

The study identified that the heart rate and pupil diameter are valid physiological measures to predict response time and quality of drivers. The study introduced PerSpeed and PerAngle as new measures of the takeover quality assessment. The findings identified a correlation between normalised mean heart rate, pupil diameter, PerSpeed and PerAngle. In addition, window size, in which mean values were calculated, had a substantial effect on the identified correlations. Such effect was reported in previous studies as explained in section 6.2. To summarise, the window size offers a focus on the overall change which abstracts the differences that happened in between. More details are furnished in sections 6.2 and 6.8.

Predicting response time and the quality of driver's performance in the takeover scenarios was crucial for highly automated driving to plan ahead to ensure the safety of drivers. There have been several papers studying different fixed transition times between 2 to 12 seconds (Gold et al., 2016, 2013a; Körber et al., 2016; Zeeb et al., 2015). With this study's findings, automation systems could identify a dynamic transition time based on the driver's physiological state to allow for the safest handover process possible.

Based upon PerSpeed and PerAngle performance measures, automation systems could predict the response time and quality of drivers during handovers even before a TOR is initiated. Using the driver's performance prediction models introduced in Chapter 5, the automation system may choose to communicate uncertainties to drivers through more alerting human-machine interfaces or resort to an emergency manoeuvre to give drivers a longer reaction time to regain situational awareness. This will, in turn, enhance the driver's experience in automated vehicles (Rakotonirainy et al., 2014).

6.2 Effect of Window Size on Physiological Measures

In this research, window size played a crucial role in the stability of the LMMs for response time, PerSpeed and PerAngle. Finding an optimal window size was commonly an essential task in similar studies (Tapia et al., 2007) and has been reported as a significant factor in the classification of EEG studies (Grimes et al., 2008). Also, a trade-off between the window size and the accuracy in many studies such as Solovey et al., (2014) were reported. Specifically, the findings of Solovey et al., (2014) showed that their longest heart rate window performed better than the smaller windows which matched with the findings in this study, see Figure 22. Essentially, a large window size meant a lag in understanding HR state (in comparison to instantaneous HR), but it provided an overall understanding of how driver's mental state affected the heart rate values.

Conversely, the findings of this study suggested that a smaller window size of PD performed better than larger window sizes. This could be because pupil diameter can change rapidly with changes in the driver's cognitive load (Klingner et al., 2011; Kramer et al., 2013). Our findings indicated that a 30 seconds window size performed well for the NDR task, response time and performance measures. Those findings match with outcomes of related studies (Klingner et al., 2011; Son et al., 2012) that concluded best performing window size was 30 seconds.

Surveyed literature used the mean value to analyse HR data (Roscoe, 1993; Solovey et al., 2014). However, mean values negate both shape and transition of the HR signal that could provide better predictability than just the mean. Further analysis was required to understand whether there is a difference in HR transitions among the groups.

6.3 Subjective Rating of the Takeovers' Difficulty

The findings indicated some interesting insight to understand some of the previously reported subjective measures of the NDR tasks, see Figure 20. The findings of this study showed an alignment with Zeeb's et al., (2016) who reported that the email task was ranked the most challenging task in a subjective rating which conflicted with their study's objective measures (e.g., deviation from lane centre). Their justifications were "email task was simply less demanding... drivers had difficulties rating their workload", (Zeeb et al., 2016).

On the contrary, the reported findings, see Chapter 4, showed that email task engagement caused a significant increase in HR during the email task followed by a significant HR peak at the TOR, see Figure 17. Thus, an email task cannot be assumed a less demanding task as Zeeb et al., (2016) reported. In fact, Salvucci and Bogunovich, (2010) reported that "Interruptions occurring at points of higher mental workload are more disruptive and lead to larger resumption lags than those occurring at points of lower mental workload". Also, writing emails had a strong correlation with stress (Marulanda-Carter and Jackson, 2012). Reflecting that on our analysis, we could assume that the email task induced a high mental workload which matches with the reported subjective measures of our study and Zeeb's et al., (2016) study. Accordingly, switching from the demanding email task to another demanding task (i.e., takeover) justified the subjective measures choices by the drivers and also reflected on their physiological behaviour as reported earlier. Another explanation could be due to a significantly degraded situational awareness (due to a higher time-off-road (Kass et al., 2007) in comparison to the TQT group that maintained eyes-on-road throughout the task.

Contrarily, subjective measures showed that the TQT was considered the least demanding task in our study by 82% of drivers. HR peak at TOR was significantly lower than the email task and so was the average HR as reported in Figure 17. During TQT, drivers spent an average of 3.3 seconds (SD=4.2) coming up with a new question. When drivers noticed a critical incident of a vehicle tailgating another on the neighbouring lane, a significant delay was reported (M=4.2, SD=1.2 seconds) among drivers to ask a new question. Such delayed responses showed that drivers performed multitasking between road monitoring and the TQT.

According to Young and Stanton, (2002), active engagement in tasks makes participants more engaged and more alert which makes it easier to takeover during TQT. Interestingly, the TQT group had a lower HR peak at TOR in comparison to the control group. Considering Young and Stanton's, (2002) findings again, the control group had no active task engagement before the TOR which turned them into a passive state causing a reduction in their level of alertness. Such an assumption could justify how the control group reported that it was difficult to engage in the takeover and why the control group had a higher HR peak than TQT group during the TOR phase. This indicates that a certain level of mental workload is preferential in the context of improved TOR quality. This should be explored in further research.

In addition to the takeover type, the order of handover processes played another role in the perception of the difficulty of the handover phase due to a previously reported learning curve of the automated system (Körber and Bengler, 2014; Larsson et al., 2014; Wright et al., 2016a). Reported subjective ratings (see Figure 20) of the takeover difficulty sorted by order showed that 40% of drivers perceived the first takeover as easy in comparison to 68% in the third takeover. Such results are taken with caution because of previously reported divergence of driving performance and subjective estimates of performance (Horrey et al., 2009b). Reported subjective measures align with the average heart rate peaks observed after TOR in the three takeovers as seen in Figure 17.

Table 14: Comparing normalised HR at each takeover

		Normalised HR mean before TOR (60-sec window)	Relative HR Peak
Takeover 1	Mean	0.47	0.126
	SD	0.125	0.22
Takeover 2	Mean	0.491	0.08
	SD	0.135	0.187
Takeover 3	Mean	0.424	0.093
	SD	0.130	0.127

To assess whether this assumption is valid, HR peak at TOR was recalculated by subtracting the mean of 'a 60 seconds normalised HR window before TOR' from the HR peak at TOR, referred to as relative HR peak in Table 14. Relative HR peak on Table 14 of the second (M=.08) and third (M=.09) takeovers were significantly smaller than the first takeover (M=.126). Relative peak HR may be considered a proxy measure of the difficulty of the takeover; especially that it matches with the reported subjective measures.

6.4 PerSpeed and PerAngle in Depth

This study introduced PerSpeed and PerAngle as two new performance measures for take-over quality in Level-3 automated driving. Minimum time to collision is an established performance measure in the field (Radlmayr et al., 2019) but it does not provide any measure of how drivers handled their vehicle. For example, a driver could employ sharp braking; this may maximise the minimum time to the collision but introduces a significant hazard to other vehicles, as would sharp transitions to a neighbouring lane. PerSpeed and PerAngle provide the missing measures that could support minimum time to collision analysis.

PerSpeed and PerAngle metrics are a further development of previously introduced performance measures such as maximum lateral and longitudinal accelerations (Gold et al., 2016), max deviation from the lane's centre (Zeeb et al., 2016), speed reduction (Larsson et al., 2014), and percent road centre (Jamson et al., 2013). The aforenoted approaches based on min, max or standard deviation were questionable to apply on nonnormally distributed variables (Urdan, 2005) such as speed or heading angle at the transition phase of this study. To avoid the limitations of the simple statistical methods, the study aimed at using an accumulative percentage change of continuous variables which gives a single value that identified the overall change introduced by drivers to speed and steering. The thesis initiated the introduction of PerSpeed and PerAngle and encouraged other studies to use as a standard for assessing performance quality to provide comparability among studies. The study identified no open access datasets that could have been used to cross-validate those measures on previous studies.

PerSpeed showed some interesting correlations with the physiological measures. Even though PerSpeed showed no statistical difference among the NDR task groups, PerSpeed correlated with the physiological measures of drivers. This could be due to some drivers who can handle secondary tasks better than others; thus, no observed increase in heart rate and pupil diameter. Other studies reported that experienced drivers would respond better than new drivers under the same secondary task condition (Ko and Ji, 2018) causing less stress and probably less increase in their physiological behaviour (Adler et al., 2000). This could explain why PerSpeed or PerAngle have no statistical significance among secondary task groups.

The fact that PerSpeed and PerAngle showed a strong correlation with driver's physiological measures highlighted the possibility that it might have a strong correlation with driver's mental workload (Brookhuis and De Waard, 2010; Marquart et al., 2015). Previous studies reported braking as a link to driver's lack of situational awareness (Zeeb et al., 2015), which shows the possibility that situational awareness could also be correlated to PerSpeed and PerAngle. Hence, future work is required for further exploration.

Finally, another finding is that this study provided experimental evidence that driver's physiological measures prior to a takeover are capable of predicting drivers' response time and quality as predicted by Chan and Singhal, (2015). It also aligns with the vision introduced by Rakotonirainy et al., (2014) that predicting driver's behaviour could enhance their experience. Based on previous suggestions, Heger et al., (2010) a mental workload recognition system using EEG and machine learning techniques is extendable to highly automated driving scenarios. Findings of this study have enabled a machine learning model to be built to accurately predict drivers' performance measures and the response quality of the takeover.

6.5 The Impact of the Learning Effect on Drivers' Performance

The results reported a decrease in both average heart rate and pupil diameter with every new takeover transition. The decrease of the heart rate data -see sections 4.2 and 4.3- could be interpreted as a decrease in the stress level caused by the uncertainties of the takeover process. As identified in a previous study, drivers enhanced their dual task skills and sped up their task switching skills (Körber et al., 2015b). The results of this study highlighted the physiological behaviour of drivers who passed through a similar experiment run by of Körber *et al.*, (2015) who identified that there was a strong relationship between the response time and the multitasking test. The relationship between multitasking skills and response time aligns with the assumptions that the main cause of the decrease of the heart rate data throughout the takeovers was due to the decrease in stress levels and observed improvement in the multitasking and task switching skills. While this study didn't collect multitasking skills test scores of its participants, participants showed varying differences recovering from the NDR tasks which could be related to their varying task handling skills. As identified earlier, there was no correlation between the driver's performance measures and the NDR tasks even though there was a strong correlation between the driver's performance and driver's physiological measures. This difference may be due to driver's multitasking and task switching skills as reported by Körber *et al.*, (2015). Finally, the results showed that driver's adaptation wasn't linear which aligns with Körber's *et al.*, (2015) findings.

6.6 Influence of Other Factors on the Driver's Performance

As identified by other studies, there were other factors affecting response time and quality such as fatigue (Driver, 2014), age (Körber et al., 2016), traffic density (Gold et al., 2016), weather conditions (Louw et al., 2016) and driving experience (Larsson et al., 2014). Those variables were not taken into consideration due to the delimitations of the study, refer to section 3.5; however, it should be considered for future work, see section 77.5.

6.8 The Effect of Window Size on the Machine Learning Algorithms

The study showed a varied number of window sizes where classification performed best. For example, the 120-second window was the best window size in which reported algorithms performed at, followed by 30 and 90 seconds, see Table 13. Such results indicate that both short and long-term observations of the drivers are essential to gain an overall understanding of their mental state. Those results align with Solovey *et al.*, (2014) findings that 30 seconds window size performed best in classifying driver's workload.

Moreover, Solovey *et al.*, (2014) tested a range of 10 to 30 seconds where 30 seconds, the largest window size performed best. The same findings were reported by [Liang, 2007] where their largest window size of 40 seconds, performed best. Finally, Wijsman *et al.*, (2011) reported efficient mental stress detection classifier using 120 seconds window. Those findings, in addition to this study's, indicate that features generated using short and long windows over driver's physiological data enrich the algorithm's input and therefore strengthen its predictability.

6.9 Feature Selection

This study explored several feature extraction methods to maximise the information gain of the classification. Figure 23 shows different groups of features based on the frequency domain and wavelet transform. In addition, other features were based on the time domain such as linear trend, autocorrelation, mass quantile and others. However, some of those features showed strong collinearity among each other which usually causes trained models to be prone to overfitting (Fulcher and Jones, 2014). When using the features subset (see Figure 23) selected by our proposed evolutionary method, a significant increase in the models' accuracy was observed.

Several coefficients in the frequency domain were reported as the top-ranked features. Such results match with excessive work done in domain-agnostic time series classification algorithms that were based on Symbolic Fourier Approximations (Schäfer and Högqvist, 2012). Those approaches were primarily based on identifying features found on the frequency domain's coefficient values. More algorithms from this family were surveyed in Bagnall's et al., (2017) review even though their investigation focused on single variate time series and missed the WEASEL algorithm (Schäfer and Leser, 2017b) because it was published later after their work.

In this study, we examined WEASEL+MUSE algorithm which is based on WEASEL's approach in identifying the top frequency domain's coefficients that maximise the information gain. It was evident that WEASEL+MUSE tended to overfit the training data. Conversely, the study's feature-based approach didn't have any overfitting issues even though it had several hyperparameters to be optimised for.

This contrast between the two approaches could be due to several reasons. Mainly, the feature selection approach used by WEASE+MUSE provided extremely effective information gain leading to overfitting. The second reason could be due to the linear classifier used by MUSE+WEASEL that may have not provided an extra layer of feature learning in contract to tree-based models that outperformed other approaches. Specifically, tree-based models of this study handled the frequency domain's coefficient as continuous variables, and by splitting the tree based on those values, the tree models provided an additional layer of learning (Breiman, 2001). However, the linear classifier in WEASEL+MUSE takes a vector of words (each word is a series of characters) which has already been estimated by the SFA algorithm. Such input to a linear classifier makes it very difficult to generalise and prone to overfitting. Finally, the reported overfitting could be due to the necessity of hyperparameter optimisation which wasn't possible due to the heavy computation of the algorithm and to provide a fair comparison with a feature-based approach that didn't receive any parameter optimisations. This aligns with other studies reporting WEASEL+MUSE as a computationally demanding algorithm (Nguyen et al., 2018).

Several coefficients in both HR and pupil diameter time series were another intriguing observation on the list of features in Figure 23. While coefficient 14 had a significant information gain, there's no real-world interpretation of its value even though it helped the model achieve reasonable accuracy. The study couldn't find a clear explanation to the meaning of this coefficient because there are no studies about interpreting HR coefficients in the frequency domain, and because the interpretation was out of the scope of this thesis. In addition, due to the limited size of the dataset, the study may not be able to provide a meaningful interpretation or conclusion to those coefficients because they may change as the dataset grows. This study acknowledges that the small dataset size was one of its research limitations that required future work.

6.10 Summary

The chapter went through the main discussion points of the thesis. The window size played a significant role in the stability of both linear and machine learning models. The correlations between the physiological behaviour of drivers and their behaviour in highly automated vehicles were discussed and related to similar studies. The study identified answers for the differences between subjective and objective measures of the driver's performance in the literature. The reasoning behind those differences based on the study's findings was explained. In summary, it was explained by the reported correlation between the driver's physiological measures and their reported subjective difficulties of the takeover scenarios.

In addition, the driver's performance measures used in the study, PerSpeed and PerAngle were discussed and correlated with similar studies. The new proposed methods were found to correlate with driver physiological data. These correlations enabled the study to build models to classify the driver performance using their physiological data. Finally, the study used a feature selection approach. This approach initiated a better understanding of the relevant features that had the highest impact on the information gain of the machine learning. In summary, the study of relevant features showed that the frequency domain of the heart rate and pupil diameter changes had the highest information gain. This understanding lead to the gain domain knowledge in understanding how physiological measures correlate with the driver's performance in the field of highly automated driving.

7

CHAPTER

Conclusion

7.1 Introduction

The focus of this study was to explore the relationship between the driver's physiological data and their behaviour in highly automated driving scenarios. The studies in this thesis didn't aim at providing a production-ready driver's mental state assessment model that could be deployed in highly automated vehicles.

However, the study provided statistical evidence of the predictability of the driver's performance during the mandatory transition period. Mainly, the study provided a better understanding of the driver's physiological data and their effect on the driver's experience and performance in highly automated vehicles. The study also answered several questions regarding the driver's learning curve and how it affects their subjective measures of the take-over difficulty. The results of the studies in this thesis give some practical evidence and suggestions about achieving a more harmonious human-machine interaction in highly automated driving between the driver (the operator) and the machine (automated driving system).

In the following section, research questions provided in chapter 1 are reviewed and answered and practical conclusions are provided.

7.2 Review of the Study's Investigations

The study was designed to answer eight questions. Questions are reported and answered in brief based on the findings of this thesis's results. To achieve the objectives of the study, this Chapter goes through the objectives and their corresponding research questions.

Objective 1: To conduct a critical review of existing literature to study different approaches assessing drivers' physiological behaviour and its effect of their performance during the 'transition' phase.

The first objective was motivated by the research question:

1. What physiological patterns that could be collected in a highly automated driving environment to provide an assessment of the driver's vigilance?

The literature of highly automated driving was very limited in studying how the physiological data could correlate with driver's vigilance. Driver's vigilance is a short term that has been defined earlier in chapter 2, but in short, it means that the driver's mental workload and situational awareness are adequate to perform the driving task. While the term is widely used in manual driving studies, it has some limited use in highly automated driving ones.

The main motivation of the study was to assess the predictability of physiological data and how it correlated with the driver's performance during the transition from automation to manual driving. The literature review of this thesis identified many studies in manual driving that achieved satisfactory results in predicting the driver's performance. Those predictions were based on physiological patterns in addition to features relating to the driver's physical positioning.

However, highly automated driving studies were not as rich as manual driving studies in that field. Specifically, no studies used a statistical approach in identifying the correct physiological data (and features) to build up a machine learning model that could predict the driver's performance during the transition phase. That's to the best of our knowledge. This was the identified as the main motivation of this research.

The literature review identified that highly automated driving studies relied heavily on eye movements, blinking, pupil diameter, eyes fixation and other physiological measures as outlined in chapter 2. Studies used the aforementioned features to assess the driver's performance, response time and mental workload during the transition. However, few studies investigated pre- and post-transition periods. In addition, heart rate data was not thoroughly studied along with pupil diameter changes throughout the automated driving scenarios. This was the main gap of the literature and acted as a starting point for the first study.

Objective 2: To design a driving scenario to assess driver's response quality during both manual and highly automated driving. The driving scenario involves the transition from highly automated to manual driving to understand driver's physiological behaviour pre- during and post-transition.

Objective 3: To conduct the study, produced on objective 3, on recruited participants.

The second and third objectives were motivated by the research question:

2. How does automation affect physiological patterns of drivers in highly automated driving scenarios?
3. How do secondary tasks reflect on driver's physiological patterns pre- during and post-transition period?

Chapter 4 presented a driving simulator study conducted to asses a driver's physiological patterns in highly automated driving scenarios. Data collection started

200 seconds before the transition and continued until the end of the experiment. Repeated measure approach was used to understand if learning is a variable in the driver's physiological behaviour.

Findings showed that drivers under the influence of a secondary task performed significantly worse than the control group who were strictly asked to maintain mental and visual vigilance to the road. Those results align with other studies (e.g. Merat *et al.*, 2012; Louw *et al.*, 2016).

During the pre-transition period, physiological patterns of drivers were different. Statistical tests showed that mean heart rate and pupil diameter are higher than the control group's mean values. The increase in heart rate and pupil diameter indicated higher mental workload which was identified using a linear model. While it was simple for the automated system to identify the driver was engaged in the email task, it becomes significantly difficult to identify mental workload that doesn't require visual distraction such as the twenty questions task or just general mind-off-road distraction (Schooler et al., 2011). However, the increase in mean heart rate, the magnitude of eye movements and mean pupil diameter could be an indicator of a driver's high mental workload.

Another interesting finding of the study was the fluctuation patterns in heart rate between the two secondary tasks. First, there was no reported statistical difference between physiological data and the type of secondary tasks. However, using a timeseries visualisation algorithm, it was evident that those timeseries fluctuated differently. The difference between the statistical approach (RM-ANOVA) and the time-series approach (Markov Transition Field) could be due to the fact that RM-ANOVA was fed the mean values of the physiological data which abstracts any fluctuations in the data. Usage of Markov Transition Field to visualise physiological data was introduced in this study and had no previous exposure in the field of Human Factors.

As the MTF approach is based on the frequency domain's observed changes, those findings opened up a new space to find appropriate features to predict the driver's performance measures. In chapter 5, most of the features that were most useful to the machine learning models were based on the coefficients of the frequency domain. These findings suggest that time-series fluctuations could have precious information that may be abstracted or obfuscated by using mean and standard deviations data to compare them. More work is required to shift the conventions of Human Factors researchers towards more modern approaches; this study doesn't provide a systematic approach to that, but it suggests that it may be required for more in-depth research.

The repeated measure design showed that the observed changes in driver's physiological mean values correlated with the driver's experience in performing the NDR task. Experienced drivers with NDR tasks (or multitasking) were found to be more relaxed performing NDR tasks because they are more familiar with them. Subsequently, those subsets of drivers performed as good as the control group.

Additionally, the repeated measure design imposed a limited time for the drivers to engage in the driving task. Although studies that prolongated their automated driving periods reported fatigue and drowsiness near the end of the experiment (Jarosch et al., 2017). While this limited time was considered a limitation of the study since drivers didn't immerse in the secondary tasks, it also helped avoid fatigue, boredom and drowsiness effect on the driver's performance.

At the start of the transition period, specifically at the TOR, a significant peak of HR was observed among drivers. The study discussed in Chapter 4 argued that it had a strong correlation with the driver's subjective measures of how difficult the take-over was. Several studies reported differences between objective measures of secondary task difficulty and the subjective measures reported by the drivers (Horrey et al., 2009b; Zeeb et al., 2016). This study's findings indicated that subjective measures aligned with the driver's perception of difficulty during the transition following the secondary task rather than their performance. Other factors that may explain this divergence could be drivers' task switching (Funke, 2007) and multitasking (Salvucci and Bogunovich, 2010) abilities. This suggestion requires more in-depth studies to assess its credibility.

There was an observed relationship between the HR peak at the TOR and the HR during the NDR task engagement in the pre-transition period. Specifically, the more engaged the driver in the secondary task (i.e, engagement measured through subjective measures and eyes off road time), the higher the stress level during the transition. Therefore, a strong peak, relative to their stress, was observed. Those findings align with (Zeeb et al., 2016) findings that some secondary tasks are more stressful than others. In addition, it provided better explanations for the strong differences between the driver's reported subjective and objective measures of the difficulty levels of the NDR tasks.

Interestingly, there was a strong correlation between the HR peak at the TOR and the order of the transitions. Specifically, the first transition had the highest peak, and its mean value went gradually down until the third transition performed by the driver. This observed learning effect was reported in various studies (e.g. Körber and Bengler, 2014; Larsson *et al.*, 2014; Wright, Samuel, Borowsky, Zilberstein and Fisher, 2016) and this study provided physiological evidence to this observation. Thus, car manufacturers need to account for the reported learning effect during the early adoption of the automated systems. The human-machine communication channels should aim at alleviating the stress level caused by the transitions to manual driving. This, in turn, will enhance driver's perception of the automated systems and encourage its adoption among other drivers.

Objective 4: To define an evaluation framework that assesses the efficiency of the prediction model produced by objective 6.

Objective 5: To assess the correlation between physiological patterns and driver's performance.

The fourth and fifth objectives were motivated by the research question:

4. What are the suitable driver's performance measures to assess their responses during the 'mandatory' transition period?
5. What's the relationship between physiological data and the driver's performance during the transition?

Most studies used Time to Collision as a performance measure (Radlmayr et al., 2019). Even though it was widely accepted in the literature in manual driving, TTC may not be suitable to be the only performance measure in highly automated driving because it omits other important variables. For example, drivers may brake aggressively to avoid a collision; hence maximising the TTC but that action may have caused a disruption in the motorway when it was possible to safely change lanes.

Due to the lack of TTC depth, every study investigated an additional performance measure that is suitable to their scenario without necessarily reflecting on others work. The reported performance measures have many similarities among them and some of those are covered in the literature review in Chapter 2. However, Radlmayr *et al.*, (2019) had an in-depth literature review of different Drivers' Performance measures. The Radlmayr's *et al.*, (2019) literature review reported that the automated driving studies had no convention in measuring or assessing the driver's performance. The researches also proposed an in-depth driving performance measure framework named TOPS that covered a wide variety of performance measure. TOPS was published after the end of the study and therefore, wasn't included in our analysis.

Prior to TOPS proposal, most of the performance measures relied heavily on the proposed scenario of their study which made them unsuitable for other studies. Consequently, this study aimed at proposing performance measures that 1) are scenario independent, 2) are easy to understand, and 3) provide quantitative values to describe the drivers' manoeuvre.

As a result, PerSpeed and PerAngle values were proposed to describe how the driver's speed and the heading angle changed during critical hazard manoeuvre in a mandatory transition period. The definitions and mathematical formulas of PerSpeed and PerAngle are laid out in chapter 4. While TOPS is a general framework that could incorporate several values, PerSpeed and PerAngle are much easier to estimate, understand and compare to. They could also be incorporated into the TOPS framework. More work is required to assess the generality of PerSpeed and PerAngle and how they compare to TOPS and other performance measures. PerSpeed particularly showed a linear relationship with heart rate and pupil diameter values. Contrarily, it showed no correlation with the type of secondary tasks.

Objective 6: To develop a system that will take data collected from objective 4 to determine the outcome of the takeover done by the driver.

The sixth objectives were motivated by the research questions:

6. What features could be extracted from physiological data that could have a predictability power of the driver's performance?

7. How could physiological data be used to assess the predictability of the driver's performance?

The study used two different features extraction methods. Both approaches generated features that were able to fit the model efficiently. However, features based on the frequency domain and wavelet had the highest information gain to the prediction algorithms. The study refrained from exploring the meaning of those features due to several reasons. First, the dataset size was small in addition to the lack of literature in the field of explaining different frequencies of the heart rate and pupil diameter data. Results showed the feasibility of predicting the driver's performance based on their physiological data during the pre-transition period. Window length played a major role in the stability of the algorithm which aligned with other studies in the manual driving domain (e.g. Erin T. Solovey *et al.*, 2014).

PerSpeed and PerAngle values showed signs of high dimensionality. That was apparent because regression models failed to predict them accurately. However, classification algorithms were able to assign drivers to high, medium and low-risk classes which is still sufficient for the automated system to draw an understanding of driver's future performance in case a mandatory transition is required. This knowledge may help the system engage the driver using a different medium or remind them to monitor the road or use a safe manoeuvre to avoid a collision.

7.3 Reflection on the Study

7.3.1 Methodology

The methodology presented in chapter 4 had some fundamental limitations. The number of participants and their age range were the main key limitations of the study. Recruited participants represented a small group of drivers' demographics which in turn could have some effect on the results of the study. However, the study's main aim was the feasibility assessment which proved to be successful. Future work will require more participants and recruitment should be extended to a larger number of participants.

In addition, the study has not considered a 3x3 ANOVA analysis to study the correlation among the secondary tasks and the order of their execution at the scenario. More participants would have been considered to achieve such results.

Future work will require different scenarios that incorporate different road scenarios. For example, most studies performed their transition on a straight road (Gold et al., 2013a; Zeeb et al., 2015) but other studies showed that take-over time might take longer in tunnels (Mok et al., 2017).

The study alerted drivers to perform a mandatory transition through an audio alert. However, there are several studies that showed other alerting methods might be better at gaining drivers' attention (Bazilinskyy et al., 2018). Those different interfaces may vary driver's stress levels, and it would be interesting to see how they may influence heart rate peak at the TOR that is reported in this study.

7.3.2 Driving Simulator

Driving simulator studies provide a significant advantage to control the virtual driving environment and provide a safe space to expose drivers to danger without risking their lives (Saffarian et al., 2012). While several studies provided evidence that simulator-based results correlate with real-life driving behaviour, there's little to no evidence that this rule still applies to highly automated driving (Louw, 2017).

Loughborough University' Design School driving simulator provided a great space to host highly automated driving studies. However, the simulator lacked the motional feedback and provided a sufficient, yet out-of-date graphics to drivers. In addition, the steering wheel posed a challenge in the realism of the driving scenario. In a real-world scenario, the steering wheel moves along with the vehicle's movements. In addition, if the driver left the wheel with an offset, this offset will still persist at the start of transition causing the drivers to put an additional effort to stabilise the vehicle. This issue was fixed by automatically notifying the drivers to centre the steering wheel at the start of automation. This simple solution helped avoid biasing the results or eliminating participants as seen in other studies (Louw, 2017).

Another significant challenge with the driving simulator was due to the limitations of their support to automated systems. This was overcome by using a combination of software commands to simulate Level-3 automation. The scenarios were cross-validated by the experimenter with other simulators and real-world driving (Tesla S model) and provided nearly the same experience.

7.3.3 Data Collection Devices

The study used two devices to collect heart rate and eye movements. Eye movements were collected using Tobii Pro Glasses. The device provided simple and portable glasses with additional infrared sensors to collect eye movements, pupil diameter and head positioning. Without a doubt, the data collection was simple using the Tobii Glasses. However, the reliability of eye blinking detection was poor and therefore excluded from the study.

HR data was collected using the Polar H7 device. The sensor was very accurate and was easy to place on males. However, it was difficult to use with females due to their brassiere's metal underwire. This was overcome by asking female participants to wear a sports bra with no metal underwire and invite a female chaperone to help them place the sensor on. The device also collected data at 1Hz which was sufficient to run the analysis but may have posed a limitation on finding more interesting patterns in heart rate variability.

7.3.4 Assessing Performance

Up until TOPS was proposed (Radlmayr et al., 2019), there were no literature reviews or a unified proposal to assess the driver's performance in highly automated driving. In this study, the main focus was correlating physiological behaviour with the driver's performance. Thus, using PerSpeed and PerAngle was adequate for the study's main goal.

Interestingly, the performance measures in this study covered the three metrics groups proposed by Radlmayr *et al.*, (2019). Specifically, vehicle guidance parameters that were covered by PerSpeed and PerAngle. In addition, subjective rating parameters which were covered by the reported perceived criticality of drivers for each take-over. Finally, mental processing parameters were considered through eye movements, pupil diameter and take-over time. A limitation of this study's approach was the lack of finding a way to merge those different parameters to give a single value at the end of the study.

7.4 Contribution to Knowledge

The main contribution to knowledge of this study is the feasibility of predicting drivers' ability to respond to critical hazards using physiological behavioural data collected before the transition. With the study's findings, automation systems could adopt a dynamic transition time based on the driver's physiological state to allow a safer transition or alert drivers who are severely out-of-the-loop using gradual warning or communicate systems' uncertainties. In addition, it provides an insight into driver's readiness and therefore, allows the automated system to adopt the correct driving strategy and plan ahead to enhance drivers' experience and make the transition phase safer for everyone.

The human factors field spent a significant effort understanding human's interactions with the automated driving systems. The field has proven that the human-machine interaction in highly automated driving environment posed several technological, legal and ethical challenges for policy makers and researchers (Louw, 2017). Most research found in this area attempts to study human's perspective on the human-machine interaction. However, the literature lacks the explorations of enabling the automated system to monitor the driver's readiness or suitability for the supervisory role over the automated driving systems. The study proved the possibility of using a machine-learning-based approach to assess human's readiness for a potential take-over and predict the riskiness level of their response to a future critical hazard. The study identified that the physiological changes of drivers were valid predictors for the machine-learning model. Finally, the work of this thesis lacked a comprehensive investigation into all the variables affecting the driver's performance. Future work is required to provide a fully functioning predictive model of the driver's performance in handling critical hazards in highly automated driving scenarios.

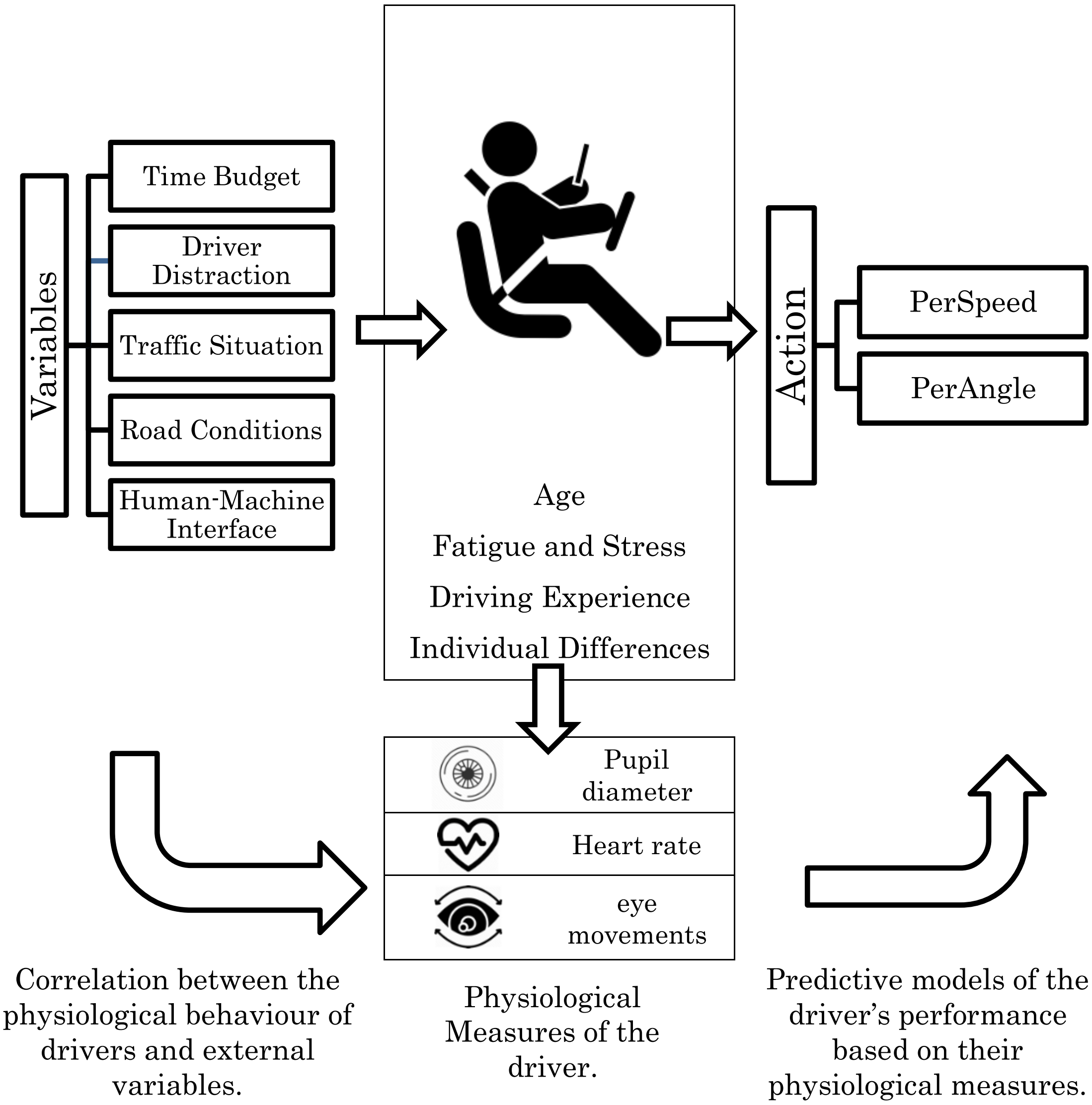


Figure 27: An illustration of the effect of external and internal variables affecting the driver's performance, their effect of driver's physiological behaviour and the performance measures introduced in this study assessing the driver's manoeuvre during the handover process.

The study has identified nine variables affecting the driver's performance during the takeover process. The variables were identified through an extensive literature review (see Chapter 2). Figure 27 lays out a framework based on the findings of this thesis considering driver background measure, vehicle and road-based variables affecting their performance. The relationship between those variables and the driver's physiological behaviour is explained briefly in section 2.4. The study has identified a relationship among the road/traffic factors, driver's physiological behaviour and their driving performance, as illustrated in Figure 27. The study used physiological measure data to build predictive models using machine learning algorithms. The results have shown strong potential in predicting the driver's performance using physiological measures of drivers. More details about the four contributions to knowledge are explained in the next sections. Variables affecting the driver's performance are discussed in section 7.4.1. Physiological changes and their cues to the deriver's performance are discussed in section 7.4.2. Performance measures of drivers are discussed in section 7.4.3 and finally predictive models of the driver's performance are discussed in section 7.4.4.

7.4.1 Variables Affecting the Driver's Performance

The study has explored all variables reported in the literature affecting the driver's performance in takeover situations when a critical hazard is present. As seen, drivers in Level-3 automated systems are decoupled from the physical control loop (Louw, 2017). When a takeover is required, the driver's ability to re-engage depends heavily on several internal and external factors. The literature review of this study identified five external and four internal factors that affect the driver's performance, see Figure 27. Two of those factors are dependent on manufacturers' choices, namely time budget and human-machine interfaces.

Road and traffic conditions are other factors affecting the driver's performance. Even though they're dynamic, (i.e., changing over time), they have shown a strong intercorrelation with both time budget and human-machine interfaces. Based on the literature review, most studies considered a pre-determined time budget and unified human-machine interface. However, manufactures should consider using adaptive time budget and varying human-machine interfaces that could adapt to driver's reaction times and help drivers re-engage more quickly. Human-machine interfaces should also be designed to direct driver's attention to the hazards to speed up the restoration of their situation awareness and communicate where potential hazards could be. This, in turn, would potentially enable drivers to respond more quickly and efficiently.

Additionally, there is a group of driver-based factors influencing their performance. The driver-based factors could be transitional or static variables. Transitional factors such as fatigue and stress are variables that could change within-session or from one driving session to another. Static factors such as age and driving experience may have a direct or indirect influence on the driver's performance at each session. For example, age and experience could indirectly influence the driver's relationship with the automated system (complacency, trust, etc.) which in turn affects their performance. Those factors change over time but will remain constant at each session. This adds time and interdependent complexities in the relationship between static and dynamic factors and the driver's performance measures such as reaction time, PerSpeed, PerAngle or eyes-off-road times.

Unfortunately, most studies assume that drivers' skills will not deteriorate over time. With the use of automation, some researchers raise the fear that drivers may lose their manual driving skills and situational awareness comprehension abilities over time (Louw, 2017). This phenomenon was observed in the flight control research literature when pilots' manual flying skills deteriorate significantly due to excessive use of autopilot systems (Casner et al., 2014).

The list of proposed factors affecting the driver's performance in this thesis lacks other subjective factors that are difficult to collect and may change over time. For example, the trust level in the automated system had a great influence on the driver's behaviour and thus influenced their performance significantly (Gold et al., 2015c).

However, the study's main hypothesis suggested that those variables may be indirectly quantified using the physiological measures of the drivers. More work is required in this matter.

Finally, drivers' individual differences in gaining situational awareness or performing physical control must be considered by manufacturers (Louw, 2017). Driver's personality or mood may have a significant impact on the driver's performance as few studies have suggested (Carsten et al., 2012; Gold and Bengler, 2014; Petermann-Stock et al., 2013). Though, there's reported evidence that support systems such (e.g. the automated driving system) may induce behavioural adaptation over time (Markkula et al., 2012). This, in turn, suggests further work on understanding driver's adaptation to the automated driving system over time. Both the thesis and all the reviewed studies in the Level-3 automated systems have not studied the adaptation of drivers to the automated system or the effect of prolonged use of the automation on their driving experience or style. This is another area that needs more exploration of the literature.

7.4.2 Physiological Changes as Cues to the Driver's Performance

Drivers' performance was the main focus of most studies in the field of highly automated driving. The surveyed studies used physiological measures to understand the driver's behaviour during the transition. Some studies used several physiological measures as objective measures of performance to understand the effect of several external or internal factors such as the driver's age or road situation, etc. The interconnected nature of external and internal factors affecting the driver's performance was the main motivation to study their effect on the driver's physiological behaviour. The main hypothesis was to assess whether physiological measures alone could be valid predictors of the driver's future performance in critical transitions.

The research has shown that physiological measures had a high correlation with the driver's performance during the transition. The manipulation of the driver's physiological measures using cognitive and visuo-cognitive NDR tasks proved the existence of the relationship between physiological measures and the driver's mental state. That relationship helped build predictive models later as explained in section 7.4.4.

The heart rate of drivers provided prominent cues to the driver's mental state and were strong covariates in clustering the driver's performance based on reaction time, PerSpeed and PerAngle, the proposed measures of performance in this study. The study showed that heart rate normalisation was essential for comparability and for building generic linear models. While normalisation is a widely accepted approach in the field of machine learning, the study aimed at introducing normalisation to the field of human factors. In addition, the study aimed at proposing timeseries visualisation approaches from the field of signal processing to the field of human factors. The visualization techniques provided subjective comparability among different NDR task groups as proposed in Chapter 4.

Pupil diameter showed a strong correlation with the driver's mental and visual workload. The results of this study showed that pupil diameter had a strong correlation with the driver's mental state and their performance in the transition. The collection of pupil diameter was challenged in the visuo-cognitive NDR task due to the brightness difference between the tablet screen and the simulator's projector. More work will be required to identify the offset of the pupil diameter caused by the mean brightness of the surrounding environment based on the recommendations of Pfleging *et al.*, (2016) model.

The study identified that each one per cent of heart rate increase corresponds to a 4.6% change in PerSpeed and each one per cent increase in pupil diameter corresponds to 9.1% change in PerSpeed. In addition, each one per cent change in heart rate corresponds to approximately 54% change in PerAngle and for each one per cent in pupil diameter, 71% change is expected. When analysing heart rate and pupil diameter, the window size played a crucial role in the stability of the aforementioned percentages. Future studies need to further investigate those percentages and their extendibility to larger datasets.

The exploration of heart rate changes answered a few questions raised in previous studies in the literature of human factors. For example, heart rate analysis in this thesis explained the deviation between the subjective and objective perception of the transition difficulty that few studies have reported. Even though a driver performed well according to the objective measures (reaction time, PerSpeed and PerAngle), the driver reported that the transition was difficult. When analysing the collected data, drivers reported that the level of difficulty was clearly correlated with the stress level during the transition and not by the difficulty of the transition itself. Logically, stress and difficulty may be highly correlated, but the uncertainty and the lack of 'transition handling' experience may have played a bigger role in inducing the stress level of drivers more than the difficulty of the situation itself. This also aligned with the observed gradual degradation of the driver's heart rate from the first to third transition performed by the same driver. This observation, in turn, aligned with others studied that reported a learning curve of driver's ability to handle the vehicle at the transition phase.

The studying of NDR tasks effect on the physiological measures of drivers showed that NDR tasks affected drivers differently. This, in turn, provided evidence that the driver's performance was not solely dependent on the NDR tasks. In fact, the driver's physiological measures could be an integral part of any predictive models of the driver's performance. That's because a physiological measure may be able to capture the individual stress levels of drivers caused by multitasking, task switching and experience in performing the NDR task. While this is not clear from the study's findings, the results open the door to study the effect of the driver's individual differences on their physiological measures.

7.4.3 Performance Measures of the Drivers

The study has identified a clear sparsity in the assessment of the driver's performance measures during the transition. Several studies used vehicle, driver or subjective-based measures of performance in addition to reaction time or take-over time as their main variables to assess the driver's performance. The main limitation to this approach was identified in several studies (Radlmayr et al., 2014; Zeeb et al., 2016) that recommended that reaction time and takeover time "did not necessarily provide a holistic view of whether drivers were prepared to resume manual control", (Louw, 2017).

In addition, using the rate of crashing may not be an adequate measure since the occurrence of crashes are rare in the literature of highly automated driving. Thus, researchers relied on minimum time to collision (minTTC) to assess near-crash scenarios (Happee et al., 2017). However, the minTTC lacks the depth of understanding the physical control of drivers, i.e., steering and braking. For example, a driver may rely on heavy braking to avoid a collision which endangers nearby vehicles but results in a high minTTC. This lack of depth in minTTC provoked researchers to find other performance measures that can define what a 'good' performance is. Most of the vehicle-related measures (see section 2.5) were an attempt to define what a 'good' performance measure is. To do so, many studies used standard deviation, minimum or maximum values of vehicle-related measures such as the positioning of the vehicle or actuators' positioning. Those approaches, unfortunately, lacked the continuity of safety outcome variables because they look for an anomaly (e.g., min or max) in continuous variables (e.g., steering) or look for mean or standard deviation of variables that don't follow a normal distribution.

Chapter 4 proposed PerSpeed and PerAngle to be the new performance measures that have several advantages over the existing performance measures. PerSpeed and PerAngle were more advantageous over the other surveyed measures in section 2.5 because they provided a continuous safety outcome of the vehicle on both lateral and longitudinal positionings. This resolves the issue of looking for anomalies in continuous variables and provides a holistic understanding of driver's physical control. PerSpeed and PerAngle could also be extended to all automated driving scenarios which nominate them to become the standard performance measures of assessing physical control of drivers in the transition. Finally, Chapter 4 identified a strong correlation between the proposed performance measures and the driver's physiological behaviour which was in turn used to build predictive models of the driver's performance in Chapter 5. Future work is required to compare various vehicle-related measures to PerSpeed and PerAngle to find their limitation and possibly merge them with other proposed techniques such as TOPS (Radlmayr et al., 2019).

7.4.4 Predictive Models of the Driver's Performance

The machine learning approach was the pillar in which the predictive models were built upon. The study investigated regression and classification to predict the driver's performance during the transition. The study used driver's physiological measures collected before the transition as an input. Results showed the possibility of predicting driver's response time, in addition, to classify them into low, medium or high-risk groups based on PerSpeed and PerAngle performance measures.

Classification approach had higher accuracy than regression models. Feature based-classifiers performed more effectively than time-series classifiers. Sliding windows and normalisation were crucial for the stability of the algorithms. The window size was one of the most important parameters to optimise into the data analysis.

Simple feature generation was sufficient to predict performance measures. Though extensive feature generation enriched the knowledge exploration of the dataset nature and proved that frequency domain-based features boosted the information gain of the classifiers. Though, there was an average increase of 2-5% in the overall accuracy of predictive models that used the relevant feature set method in comparison to the models that used the simple feature generation method. A drawback to the relevant feature set method was the complexity of the method, the vast number of parameters that need to be tuned and finally the processing time. A deep learning approach should be used in the future should a large dataset be available. Out of all used classifiers, Random Forest classifier outperformed in predicting all the driver's performance measures. This is a typical finding and aligns with other studies.

7.5 Future Work

The research has initiated the exploration of assessing the predictability of the driver's performance using the driver's physiological measures. Through the literature review, the study identified 9 different factors affecting the performance of drivers as stated in section 2.3. The driver distraction was the only variable manipulated on the driving simulator study. The thesis has explored the delimitations of the study in section 3.5 and listed the other factors that need to be explored for future work. Future work needs to consider a 3x3 ANOVA analysis to assess the correlation between the order of takeovers and the NDR tasks.

The feature analysis of this study identified several frequency-domain-based features that enriched the machine learning models. The understanding of those frequencies and their values were very limited. More research will be required to find an interpretation of different frequencies of the heart rate and pupil diameter and how they could possibly link to driver's mental transitions or state. The literature has very limited resources on that matter, and future work is suggested here.

Further work is required in building up and testing a dynamic human-machine interface that incorporated the prediction of the proposed models in Chapter 5. The suggested experiment should explore the effect of static and dynamic human-machine interfaces. The dynamic human-machine interface could alert drivers which are expected to perform poorly whenever the predictive models suggested a potential poor performance to alert drivers to get back in the loop.

The limited resources of the academic study made it difficult to recruit a large number of participants. A suggestion for future work is to expand the data collection to a larger number of participants while considering other factors affecting the driver's performance such as road or traffic conditions. Thus, deep learning approach could be applicable to be used. Finally, based on the findings of the study, the researcher suggests the exploration of other physiological methods such as head pose and heart rate variability.

7.6 Summary

The research conducted in this study listed a set of factors affecting the driver's performance in highly automated vehicles. The study proved that physiological measures were able to provide a better understanding of the driver's mental state and provided an underdstanding to their subjective and objective measures of performance. Physiological measures were also valid features to predict the driver's performance measures using feature-based machine learning models.

Bibliography

Abbink, D.A., Mulder, M., Boer, E.R., 2012. Haptic shared control: smoothly shifting control authority? Cogn. Technol. Work 14, 19–28. https://doi.org/10.1007/s10111-011-0192-5

Adler, N.E., Epel, E.S., Castellazzo, G., Ickovics, J.R., 2000. Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. Heal. Psychol. 19, 586–592. https://doi.org/10.1037/0278-6133.19.6.586

Alford, C.A., 2009. Sleepiness, countermeasures and the risk of motor vehicle accidents, in: Drugs, Driving and Traffic Safety. Springer, pp. 207–232.

Alpaydin, E., 2014. Introduction to machine learning.

Anderson, J., Kalra, N., Stanley, K., Sorensen, P., Samaras, C., Oluwatola, O., 2016. Autonomous Vehicle Technology: A Guide for Policymakers. https://doi.org/10.7249/RR443-2

Anderson, T.W., 2011. The Statistical Analysis of Time Series. John Wiley & Sons.

Anderson, W.M., Hirshkowitz, M., Davila, D.G., Kramer, M., Johnson, S.F., Berry, R.B., Littner, M., Kapen, S., Wise, M., Bailey, D., Loube, D., Kushida, C.A., 2017. Practice Parameters for the Role of Actigraphy in the Study of Sleep and Circadian Rhythms: An Update for 2002. Sleep 26, 337–341. https://doi.org/10.1093/sleep/26.3.337

Anstey, K.J., Wood, J., Lord, S., Walker, J.G., 2005. Cognitive, sensory and physical factors enabling driving safety in older adults. Clin. Psychol. Rev. 25, 45–65.

Arendt, H., Canovan, M., 1998. The human condition. University of Chicago Press.

Armand, A., 2016. Situation understanding and risk assessment framework for preventive driver assistance.

Bagnall, A., Lines, J., Bostrom, A., Large, J., Keogh, E., 2017. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data Min. Knowl. Discov. 31, 606–660. https://doi.org/10.1007/s10618-016-0483-9

Bainbridge, L., 1983. Ironies of Automation * 19.

Baliga, R.R., Eagle, K.A., 2008. Practical Cardiology: Evaluation and Treatment of Common Cardiovascular Disorders 770.

Banks, S., Catcheside, P., Lack, L., Grunstein, R.R., McEvoy, R.D., 2004. Low levels of alcohol impair driving simulator performance and reduce perception of crash risk in partially sleep deprived subjects. Sleep 27, 1063–1067. https://doi.org/10.1093/sleep/27.6.1063

Barr, L., Popkin, S., Howarth, H., 2009. An evaluation of emerging driver fatigue detection measures and technologies.

Batmaz, I., Ozturk, M., 2008. Using Pupil Diameter Changes for Measuring Mental Workload under Mental Processing. J. Appl. Sci. 8, 68–76. https://doi.org/10.3923/jas.2008.68.76

Baumann, M.R.K., Rösler, D., Krems, J.F., 2007. Situation awareness and secondary task performance while driving, in: International Conference on Engineering Psychology and Cognitive Ergonomics. pp. 256–263.

Bazilinskyy, P., Petermeijer, S.M., Petrovych, V., Dodou, D., de Winter, J.C.F., 2018. Take-over requests in highly automated driving: A crowdsourcing survey on auditory, vibrotactile, and visual displays. Transp. Res. Part F Traffic Psychol. Behav. 56, 82–98. https://doi.org/10.1016/J.TRF.2018.04.001

Beller, J., Heesen, M., Vollrath, M., 2013a. Improving the driver--automation interaction an approach using automation uncertainty. Hum. Factors J. Hum.

Factors Ergon. Soc. 55, 1130–1141.

Beller, J., Heesen, M., Vollrath, M., Universität, T., 2013b. Improving the Driver – Automation Interaction : An Approach Using Automation Uncertainty. https://doi.org/10.1177/0018720813482327

Bennett, K.P., Campbell, C., 2000. Support vector machines: hype or hallelujah? SIGKDD Explor. Newsl. 2, 1–13. https://doi.org/10.1145/380995.380999

Berk, R.A., 2006. An introduction to ensemble methods for data analysis. Sociol. Methods Res. 34, 263–295. https://doi.org/10.1177/0049124105283119

Bibby, K.S., Margulies, F., Rijnsdorp, J.E., Withers, R.M.J., Makarov, I.M., Rijnsdorp, J.E., 1975. Man's Role in Control Systems. IFAC Proc. Vol. 8, 664–683. https://doi.org/10.1016/S1474-6670(17)67612-2

Blanco, M., Biever, W.J., Gallagher, J.P., Dingus, T.A., 2006. The impact of secondary task cognitive processing demand on driving performance. Accid. Anal. Prev. 38, 895–906.

Bramer, M., 2016. Principles of Data Mining, Undergraduate Topics in Computer Science. Springer London, London. https://doi.org/10.1007/978-1-4471-7307-6

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

Breiman, L., 1996. Bias, variance, and arcing classifier.

Brookhuis, K.A., de Waard, D., 2001. ASSESSMENT OF DRIVERS'WORKLOAD: PERFORMANCE AND SUBJECTIVE AND PHYSIOLOGICAL INDEXES. Stress. Workload fatigue.

Brookhuis, K.A., De Waard, D., 2010. Monitoring drivers' mental workload in driving simulators using physiological measures. Accid. Anal. Prev. 42, 898–903. https://doi.org/10.1016/j.aap.2009.06.001

Bruce Mehler, B.R.&, Jeffery A. Dusek, Mehler, B., Reimer, B., Dusek, J.A., 2011. MIT AgeLab delayed digit recall task (n-back). Cambridge, MA Massachusetts

Inst. Technol.

Bryman, A., 2003. Quantity and quality in social research.

Buehler, M., Iagnemma, K., Singh, S., 2009. The DARPA urban challenge: autonomous vehicles in city traffic. springer.

Buehler, M., Iagnemma, K., Singh, S., 2007. The 2005 DARPA grand challenge: the great robot race.

Bühlmann, P., Hothorn, T., 2007. Boosting Algorithms: Regularization, Prediction and Model Fitting. Stat. Sci. 22, 477–505. https://doi.org/10.1214/07-STS242

Caffier, P.P., Erdmann, U., Ullsperger, P., 2003. Experimental evaluation of eye-blink parameters as a drowsiness measure. Eur. J. Appl. Physiol. 89, 319–325.

Cain, B., 2007. A Review of the Mental Workload Literature.

Campbell, J.L., Brown, J.L., Graving, J.S., Richard, C.M., Lichty, M.G., Bacon, L.P., Morgan, J.F., Sanquist, T., 2018. Human factors design principles for level 2 and level 3 automated driving concepts 122.

Campbell, J.L., Richard, C.M., Brown, J.L., McCallum, M., 2007. Crash warning system interfaces: Human Factors insight and lessons learned. US Dep. Transp. Natl. Highw. Traffic Saf. Adm. No. HS-810, 1–179. https://doi.org/HS 810 697

Cantin, V., Lavallière, M., Simoneau, M., Teasdale, N., 2009. Mental workload when driving in a simulator: Effects of age and driving complexity. Accid. Anal. Prev. 41, 763–771.

Carsten, O., Lai, F.C.H., Barnard, Y., Jamson, A.H., Merat, N., 2012. Control Task Substitution in Semiautomated Driving. Hum. Factors J. Hum. Factors Ergon. Soc. 54, 747–761. https://doi.org/10.1177/0018720812460246

Caruana, R., Caruana, R., Niculescu-Mizil, A., 2005. An Empirical Comparison of Supervised Learning Algorithms Using Different Performance Metrics. PROC. 23 RD INTL. CONF. Mach. Learn. (ICML'06 161--168.

Casner, S.M., Geven, R.W., Recker, M.P., Schooler, J.W., 2014. The retention of manual flying skills in the automated cockpit. Hum. Factors 56, 1506–1516. https://doi.org/10.1177/0018720814535628

Cawley, G.C., Talbot, N.L.C., 2004. Fast exact leave-one-out cross-validation of sparse least-squares support vector machines. Neural Networks 17, 1467–1475. https://doi.org/10.1016/j.neunet.2004.07.002

Chan, M., Singhal, A., 2015. Emotion matters : Implications for distracted driving. Saf. Sci. 72, 302–309. https://doi.org/10.1016/j.ssci.2014.10.002

Choudhary, P., Velaga, N.R., 2017. Modelling driver distraction effects due to mobile phone use on reaction time. Transp. Res. Part C Emerg. Technol. 77, 351–365. https://doi.org/10.1016/J.TRC.2017.02.007

Christ, M., Braun, N., Neuffer, J., Kempa-Liehr, A.W., 2018. Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). Neurocomputing 307, 72–77. https://doi.org/10.1016/J.NEUCOM.2018.03.067

Christ, M., Kempa-Liehr, A.W., Feindt, M., 2016. Distributed and parallel time series feature extraction for industrial big data applications.

Cohen, L., Manion, L., Morrison, K., 2013. Research methods in education. Routledge.

Cornford, T., Smithson, S., 1996. Research approaches, in: Project Research in Information Systems. Springer, pp. 32–54.

Cover, T., Hart, P., 1967. Nearest neighbor pattern classification. IEEE Trans. Inf. Theory 13, 21–27. https://doi.org/10.1109/TIT.1967.1053964

Craye, C., Karray, F., 2015. Driver distraction detection and recognition using RGB-D sensor. arXiv Prepr. arXiv1502.00250.

Crotty, M., 1998. The foundations of social research: Meaning and perspective in the research process. Sage.

Cummings, M.L., Guerlain, S., 2007. Developing Operator Capacity Estimates for Supervisory Control of Autonomous Vehicles. Hum. Factors J. Hum. Factors Ergon. Soc. 49, 1–15. https://doi.org/10.1518/001872007779598109

da Silva, P.N., Plastino, A., Freitas, A.A., 2018. A Novel Genetic Algorithm for Feature Selection in Hierarchical Feature Spaces, in: Proceedings of the 2018 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, Philadelphia, PA, pp. 738–746. https://doi.org/10.1137/1.9781611975321.83

Dalgleish, T., Williams, J.M.G.., Golden, A.-M.J., Perkins, N., Barrett, L.F., Barnard, P.J., Au Yeung, C., Murphy, V., Elward, R., Tchanturia, K., Watkins, E., 2007. Feature engineering for machine learning and data analytics, Journal of Experimental Psychology: General.

Damböck, D., Farid, M., Tönert, L., Bengler, K., 2012. Übernahmezeiten beim hochautomatisierten Fahren. 5. Tagung Fahrerassistenz.

Darzi, A., Gaweesh, S.M., Ahmed, M.M., Novak, D., 2018. Identifying the Causes of Drivers' Hazardous States Using Driver Characteristics, Vehicle Kinematics, and Physiological Measurements. Front. Neurosci. 12, 568. https://doi.org/10.3389/fnins.2018.00568

Dasgupta, A., George, A., Happy, S.L., Routray, A., 2013. A Vision-Based System for Monitoring the Loss of Attention in Automotive Drivers. IEEE Trans. Intell. Transp. Syst. 14, 1825–1838. https://doi.org/10.1109/TITS.2013.2271052

Davis, J., Goadrich, M., 2006. The relationship between Precision-Recall and ROC curves, in: Proceedings of the 23rd International Conference on Machine Learning - ICML '06. ACM Press, New York, New York, USA, pp. 233–240. https://doi.org/10.1145/1143844.1143874

De Rosario, H., Solaz, J.S., Rodríguez, N., Bergasa, L.M., 2010. Controlled inducement and measurement of drowsiness in a driving simulator. IET Intell. Transp. Syst. 4, 280. https://doi.org/10.1049/iet-its.2009.0110

de Waard, D., van der Hulst, M., Hoedemaeker, M., Brookhuis, K.A., 1999. Driver Behavior in an Emergency Situation in the Automated Highway System. Transp. Hum. Factors 1, 67–82. https://doi.org/10.1207/sthf0101_7

De Winter, J., Happee, P., 2012. Advantages and disadvantages of driving simulators: a discussion, in: Proceedings of Measuring Behavior. pp. 47–50.

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. Transp. Res. Part F Traffic Psychol. Behav. https://doi.org/10.1016/j.trf.2014.06.016

Denzin, N.K., Lincoln, Y.S., 2005. The SAGE handbook of qualitative research. Sage Publications.

Dinges, D.F., Grace, R., 1998. PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. US Dep. Transp. Fed. Highw. Adm. Publ. Number FHWA-MCRT-98-006.

Dong, Y., Zhang, Y., Yue, J., Hu, Z., 2015. Comparison of random forest, random ferns and support vector machine for eye state classification. Multimed. Tools Appl. https://doi.org/10.1007/s11042-015-2635-0

Dongen, H.P.A. Van, 2004. Comparison of mathematical model predictions to experimental data of fatigue and performance. Aviat. Sp. Environ. Med. 75, No. 3, A15–A36.

Driver, E.T.H.E., 2014. Fatigue in the Automated Vehicle: Do Games and Conversation Distract or Energize the Driver? 2053–2057.

Endsley, M.R., 2017. Autonomous Driving Systems: A Preliminary Naturalistic Study of the Tesla Model S. J. Cogn. Eng. Decis. Mak. 11, 225–238.

https://doi.org/10.1177/1555343417695197

Endsley, M.R., 1995. Toward a Theory of Situation Awareness in Dynamic Systems 37, 32–64.

Endsley, M.R., Kiris, E.O., 1995. The out-of-the-loop performance problem and level of control in automation. Hum. Factors J. Hum. Factors Ergon. Soc. 37, 381–394.

Endsley, M.R., Onal, E., Problems, A.A., 1997. The Impact of Intermediate Levels of Automation on Situation Awareness and Performance in Dynamic Control Systems.

Eriksson, A., Banks, V.A., Stanton, N.A., 2017. Transition to manual: Comparing simulator with on-road control transitions. Accid. Anal. Prev. 102, 227–234. https://doi.org/10.1016/j.aap.2017.03.011

Eriksson, A., Stanton, N.A., 2017. Takeover Time in Highly Automated Vehicles: Noncritical Transitions to and From Manual Control. Hum. Factors J. Hum. Factors Ergon. Soc. 59, 689–705. https://doi.org/10.1177/0018720816685832

faceLAB, 2016. faceLAB 5 | Seeing Machines. Seeing Mach. RSS.

Fagnant, D., Kockelman, K.M., 2013. Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations.

Fernández, A., Usamentiaga, R., Carús, J., Casado, R., Fernández, A., Usamentiaga, R., Carús, J.L., Casado, R., 2016. Driver Distraction Using Visual-Based Sensors and Algorithms. Sensors 16, 1805. https://doi.org/10.3390/s16111805

Finley, T., Joachims, T., 2008. Supervised k-means clustering.

Fitch, G.M., Soccolich, S.A., Guo, F., McClafferty, J., Fang, Y., Olson, R.L., Perez, M.A., Hanowski, R.J., Hankey, J.M., Dingus, T.A., 2013. The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk.

Fleiter, J.J., Lennon, A., Watson, B., 2010. How do other people influence your driving speed? Exploring the 'who'and the 'how'of social influences on speeding from a qualitative perspective. Transp. Res. part F traffic Psychol. Behav. 13, 49–62.

Flemisch, F., Heesen, M., Hesse, T., Kelsch, J., Schieben, A., Beller, J., 2012. Towards a dynamic balance between humans and automation : authority , ability , responsibility and control in shared and cooperative control situations 3–18. https://doi.org/10.1007/s10111-011-0191-6

Flemisch, F.O., Kelsch, J., Loper, C., Schieben, A., Schindler, J., Matthias, H., 2008. Cooperative Control and Active Interfaces for Vehicle Assistance and Automation. FISITA World Automot. Congr. 301–310. https://doi.org/10.1016/j.compstruct.2004.08.017

Freidman, J.H., 2008. Greedy Function Approximation : A Gradient Boosting Machine. Institue Math. Stat. 29, 1189–1232. https://doi.org/10.1017/CBO9781107415324.004

Friedman, J.H., 2001. Greedy Function Approximation: A Gradient Boosting Machine. Source Ann. Stat. 29, 1189–1232.

Fulcher, B.D., Jones, N.S., 2014. Highly Comparative Feature-Based Time-Series Classification. IEEE Trans. Knowl. Data Eng. 26, 3026–3037. https://doi.org/10.1109/TKDE.2014.2316504

Funke, G., 2007. The effects of automation and workload on driver performance, subjective workload, and mood.

Gable, T.M., Kun, A.L., Walker, B.N., Winton, R.J., 2015. Comparing heart rate and pupil size as objective measures of workload in the driving context, in: Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15. ACM Press, New

York, New York, USA, pp. 20–25. https://doi.org/10.1145/2809730.2809745

Gold, C., Bengler, K., 2014. Influence of Automated Brake Application on Take-Over Situations in Highly Automated Driving Scenarios. FISITA 2014 World Automot. Congr.

Gold, C., Berisha, I., Bengler, K., 2015a. Utilization of drivetime - Performing non-driving related tasks while driving highly automated. Proc. Hum. Factors Ergon. Soc. 2015–Janua, 1666–1670. https://doi.org/10.1177/1541931215591360

Gold, C., Berisha, I., Bengler, K., 2015b. Utilization of Drivetime – Performing Non-Driving Related Tasks While Driving Highly Automated 1666–1670.

Gold, C., Damböck, D., Lorenz, L., Bengler, K., 2013a. "Take over!" How long does it take to get the driver back into the loop?, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. pp. 1938–1942.

Gold, C., Körber, M., Hohenberger, C., Lechner, D., Bengler, K., 2015c. ScienceDirect Trust in automation – Before and after the experience of take-over scenarios in a highly automated vehicle 00, 372–379.

Gold, C., Körber, M., Lechner, D., Bengler, K., 2016. Taking over Control from Highly Automated Vehicles in Complex Traffic Situations. Hum. Factors 58, 642–652. https://doi.org/10.1177/0018720816634226

Gold, C., Lorenz, L., Damböck, D., Bengler, K., 2013b. Partially Automated Driving as a Fallback Level of High Automation. 6. Tagung Fahrerassistenzsysteme. Der Weg zum Autom. Fahr. 28, 2013.

Gopher, D., Braune, R., 1984. On the psychophysics of workload: Why bother with subjective measures? Hum. Factors 26, 519–532. https://doi.org/10.1177/001872088402600504

Grimes, D., Tan, D.S., Hudson, S.E., Shenoy, P., Rao, R.P.N., 2008. Feasibility and pragmatics of classifying working memory load with an electroencephalograph.

Proceeding twenty-sixth Annu. CHI Conf. Hum. factors Comput. Syst. - CHI ’08 835. https://doi.org/10.1145/1357054.1357187

Guyon, I., Elisseeff, A., 2006. An Introduction to Feature Extraction, in: Feature Extraction. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–25. https://doi.org/10.1007/978-3-540-35488-8_1

Guyon, I., Elisseeff, A., 2003. AnIntroductionToVariableAndFeautureSelection.pdf [WWW Document]. J. Mach. Learn. Res. https://doi.org/10.1016/j.aca.2011.07.027

Guyon, I., Gunn, S., Nikravesh, M., Zadeh, L.A., 2006. Feature Extraction: Foundations and Applications (Studies in Fuzziness and Soft Computing), Soft Computing. Springer-Verlag.

Haigney, D., Westerman, S.J., 2001. Mobile (cellular) phone use and driving: A critical review of research methodology. Ergonomics 44, 132–143.

Hancock, P.A., Billings, D.R., Schaefer, K.E., Florida, C., Chen, J.Y.C., Visser, E.J. De, Parasuraman, R., 2009. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. https://doi.org/10.1177/0018720811417254

Hansheng, W., Guodong, L., Guohua, J., 2007. Robust Regression Shrinkage and Consistent Variable Selection Through the LAD-Lasso. J. Bus. Econ. Stat. 25, 347--355. https://doi.org/10.1198/00

Happee, R., Gold, C., Radlmayr, J., Hergeth, S., Bengler, K., 2017. Take-over performance in evasive manoeuvres. Accid. Anal. Prev. 106, 211–222. https://doi.org/10.1016/j.aap.2017.04.017

Hart, S.G., 2012. Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 50, 904–908. https://doi.org/10.1177/154193120605000909

Hartley, L., Horberry, T., Mabbott, N., Krueger, G.P., 2000. Review of fatigue detection and prediction technologies. Natl. Road Transp. Comm.

Harvey, C., Stanton, N.A., Pickering, C.A., Mcdonald, M., Zheng, P., 2011. A usability evaluation toolkit for In-Vehicle Information Systems (IVISs). Appl. Ergon. 42, 563–574. https://doi.org/10.1016/j.apergo.2010.09.013

Healey, J., Wang, C., Dopfer, A., Yu, C., 2012. M2M Gossip : Why Might We Want Cars to Talk About Us ? 265–268.

Healey, J.A., Picard, R.W., 2005. Detecting stress during real-world driving tasks using physiological sensors. IEEE Trans. Intell. Transp. Syst. 6, 156–166. https://doi.org/10.1109/TITS.2005.848368

Hecht, T., Feldhütter, A., Radlmayr, J., Nakano, Y., Miki, Y., Henle, C., Bengler, K., 2019. A Review of Driver State Monitoring Systems in the Context of Automated Driving. pp. 398–408. https://doi.org/10.1007/978-3-319-96074-6_43

Heenan, A., Herdman, C.M., Brown, M.S., Robert, N., 2014. Effects of conversation on situation awareness and working memory in simulated driving. Hum. Factors J. Hum. Factors Ergon. Soc. 0018720813519265.

Heger, D., Putze, F., Schultz, T., 2010. Online Workload Recognition from EEG Data during Cognitive Tests and Human-Machine Interaction. Ki 2010 Adv. Artif. Intell. 6359, 410–417.

Helldin, T., Falkman, G., Riveiro, M., Davidsson, S., 2013. Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving, in: Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13. ACM Press, New York, New York, USA, pp. 210–217. https://doi.org/10.1145/2516540.2516554

Hollnagel, E., Woods, D.D., 2005. Joint Cognitive Systems: Foundations of Cognitive Systems Engineering (Google eBook). Taylor & Francis.

Horberry, T., Anderson, J., Regan, M.A., Triggs, T.J., Brown, J., 2006. Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. Accid. Anal. Prev. 38, 185–191. https://doi.org/10.1016/j.aap.2005.09.007

Horrey, W.J., Lesch, M.F., Garabet, A., 2009a. Dissociation between driving performance and drivers' subjective estimates of performance and workload in dual-task conditions. J. Safety Res. 40, 7–12. https://doi.org/10.1016/j.jsr.2008.10.011

Horrey, W.J., Lesch, M.F., Garabet, A., 2009b. Dissociation between driving performance and drivers' subjective estimates of performance and workload in dual-task conditions. J. Safety Res. 40, 7–12. https://doi.org/10.1016/J.JSR.2008.10.011

Hs, D.O.T., 2014. Human Factors Evaluation of Level 2 And Level 3 Automated Driving Concepts Concepts of Operation.

Hu, J., Min, J., 2018. Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model. Cogn. Neurodyn. 12, 431–440. https://doi.org/10.1007/s11571-018-9485-1

Ibañez-Guzmán, J., Laugier, C., Yoder, J.-D., Thrun, S., 2012. Autonomous Driving: Context and State-of-the-Art, in: Handbook of Intelligent Vehicles. Springer London, London, pp. 1271–1310. https://doi.org/10.1007/978-0-85729-085-4_50

Ibrahim, M.M., 2014. Video processing analysis for non-invasive fatigue detection and quantification. University of Strathclyde.

Inagaki, T., Sheridan, T.B., 2018. A critique of the SAE conditional driving automation definition, and analyses of options for improvement. Cogn. Technol. Work 1–10. https://doi.org/10.1007/s10111-018-0471-5

Ingle, S., Phute, M., 2016. Tesla Autopilot : Semi Autonomous Driving, an Uptick for

Future Autonomy. Int. Res. J. Eng. Technol. 2395–56.

Jafarnejad, S., Castignani, G., Engel, T., 2018. Non-intrusive Distracted Driving Detection Based on Driving Sensing Data.

Jain, A.K., 2010. Data clustering: 50 years beyond K-means. Pattern Recognit. Lett. 31, 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

Jamson, A.H., Merat, N., 2005. Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving. Transp. Res. Part F Traffic Psychol. Behav. 8, 79–96.

Jamson, A.H., Merat, N., Carsten, O.M.J., Lai, F.C.H., 2013. Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. Transp. Res. part C Emerg. Technol. 30, 116–125. https://doi.org/10.1016/j.trc.2013.02.008

Jamson, A.H., Westerman, S.J., J. Hockey, G.R., Carsten, O.M.J., Hockey, G.R.J., Carsten, O.M.J., 2004. Speech-Based E-Mail and Driver Behavior: Effects of an In-Vehicle Message System Interface. Hum. Factors 46, 625–639. https://doi.org/10.1518/hfes.46.4.625.56814

Jarosch, O., Kuhnt, M., Paradies, S., Bengler, K., 2017. It's Out of Our Hands Now! Effects of Non-Driving Related Tasks During Highly Automated Driving on Drivers' Fatigue.

Kaptein, N., Theeuwes, J., Van Der Horst, R., 1996. Driving Simulator Validity: Some Considerations. Transp. Res. Rec. J. Transp. Res. Board 1550, 30–36. https://doi.org/10.3141/1550-05

Karvonen, J., Vuorimaa, T., 1988. Heart Rate and Exercise Intensity During Sports Activities: Practical Application. Sport. Med. An Int. J. Appl. Med. Sci. Sport Exerc. https://doi.org/10.2165/00007256-198805050-00002

Kass, S.J., Cole, K.S., Stanny, C.J., 2007. Effects of distraction and experience on situation awareness and simulated driving. Transp. Res. Part F Traffic Psychol. Behav. 10, 321–329. https://doi.org/10.1016/j.trf.2006.12.002

Kirchner, W.K., 1958. Age differences in short-term retention of rapidly changing information. J. Exp. Psychol. 55, 352.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J., others, 2006. The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data.

Kléma, J., Kout, J., Vejmelka, M., 2004. Predictive System for Multivariate Time Series. labe.felk.cvut.cz.

Klingner, J., Tversky, B., Hanrahan, P., 2011. Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. Psychophysiology 48, 323–332. https://doi.org/10.1111/j.1469-8986.2010.01069.x

Ko, S.M., Ji, Y.G., 2018. How we can measure the non-driving-task engagement in automated driving: Comparing flow experience and workload. Appl. Ergon. 67, 237–245. https://doi.org/10.1016/j.apergo.2017.10.009

Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L., Nass, C., 2015. Why did my car just do that ? Explaining semi-autonomous driving actions to improve driver understanding , trust , and performance. Int. J. Interact. Des. Manuf. 269–275. https://doi.org/10.1007/s12008-014-0227-2

Körber, M., Bengler, K., 2014. Potential individual differences regarding automation effects in automated driving, in: Proceedings of the XV International Conference on Human Computer Interaction. p. 22.

Körber, M., Cingel, A., Zimmermann, M., Bengler, K., 2015a. Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. Procedia Manuf. 3, 2403–2409. https://doi.org/10.1016/j.promfg.2015.07.499

Körber, M., Gold, C., Lechner, D., Bengler, K., 2016. The influence of age on the take-over of vehicle control in highly automated driving. Transp. Res. part F traffic Psychol. Behav. 39, 19–32.

Körber, M., Weißgerber, T., Blaschke, C., Farid, M., Kalb, L., 2015b. Prediction of take-over time in highly automated driving by two psychometric tests. DYNA 82, 195–201. https://doi.org/10.15446/dyna.v82n193.53496

Kotsiantis, S., Kanellopoulos, D., 2006. Discretization Techniques : A recent survey. GESTS Int. Trans. Comput. Sci. Eng. 32, 47–58. https://doi.org/10.1016/B978-044452781-3/50006-2

Kotsiantis, S.B., Kanellopoulos, D., Pintelas, P.E., 2006. Data preprocessing for supervised leaning. Int. J. Comput. Sci. 1, 111–117.

Kraemer, S., Danker-Hopfe, H., Dorn, H., Schmidt, A., Ehlert, I., Herrmann, W.M., 2000. Time-of-day variations of indicators of attention: performance, physiologic parameters, and self-assessment of sleepiness. Biol. Psychiatry 48, 1069–1080. https://doi.org/10.1016/S0006-3223(00)00908-2

Kramer, S.E., Lorens, A., Coninx, F., Zekveld, A.A., Piotrowska, A., Skarzynski, H., 2013. Processing load during listening: The influence of task characteristics on the pupil response. Lang. Cogn. Process. 28, 426–442. https://doi.org/10.1080/01690965.2011.642267

Kray, J., Eber, J., Lindenberger, U., 2004. Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. Acta Psychol. (Amst). 115, 143–165.

Kudo, M., Sklansky, J., 2000. Comparison of algorithms that select features for pattern classifiers. Pattern Recognit. 33, 25–41. https://doi.org/10.1016/S0031-3203(99)00041-2

Kunze, A., Summerskill, S.J., Marshall, R., Filtness, A.J., 2018a. Automation Transparency: Implications of Uncertainty Communication for Human-

Automation Interaction and Interfaces. Ergonomics 1–22. https://doi.org/10.1080/00140139.2018.1547842

Kunze, A., Summerskill, S.J., Marshall, R., Filtness, A.J., 2018b. Evaluation of Variables for the Communication of Uncertainties Using Peripheral Awareness Displays, in: Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '18. ACM Press, New York, New York, USA, pp. 147–153. https://doi.org/10.1145/3239092.3265958

Lal, S.K.L., Craig, A., 2002. Driver fatigue: electroencephalography and psychological assessment. Psychophysiology 39, 313–321.

Lal, S.K.L., Craig, A., 2001. A critical review of the psychophysiology of driver fatigue. Biol. Psychol. 55, 173–194. https://doi.org/10.1016/S0301-0511(00)00085-5

Larsson, A.F.L., Kircher, K., Andersson, J., Hultgren, J.A., 2014. Learning from experience: Familiarity with ACC and responding to a cut-in situation in automated driving. Transp. Res. part F traffic Psychol. Behav. 27, 229–237. https://doi.org/10.1016/j.trf.2014.05.008

Le, A.S., Aoki, H., Murase, F., Ishida, K., 2018. A Novel Method for Classifying Driver Mental Workload Under Naturalistic Conditions With Information From Near-Infrared Spectroscopy. Front. Hum. Neurosci. 12, 431. https://doi.org/10.3389/fnhum.2018.00431

Lee, B.G., Park, J.-H.H., Pu, C.C., Chung, W.-Y.Y., 2016. Smartwatch-Based Driver Vigilance Indicator With Kernel-Fuzzy-C-Means-Wavelet Method. IEEE Sens. J. 16, 242–253. https://doi.org/10.1109/JSEN.2015.2475638

Lee, J.D., Caven, B., Haake, S., Brown, T.L., 2001. Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. Hum. Factors J. Hum. Factors Ergon. Soc. 43, 631–640.

Lee, J.D., See, K.A., 2009. Trust in Automation: Designing for Appropriate Reliance. Hum. Factors J. Hum. Factors Ergon. Soc. 46, 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Li-Wei Ko, Wei-Kai Lai, Wei-Gang Liang, Chun-Hsiang Chuang, Shao-Wei Lu, Yi-Chen Lu, Tien-Yang Hsiung, Hsu-Hsuan Wu, Chin-Teng Lin, 2015. Single channel wireless EEG device for real-time fatigue level detection, in: 2015 International Joint Conference on Neural Networks (IJCNN). IEEE, pp. 1–5. https://doi.org/10.1109/IJCNN.2015.7280817

Li, N., Busso, C., 2015. Predicting Perceived Visual and Cognitive Distractions of Drivers With Multimodal Features. IEEE Trans. Intell. Transp. Syst. 16, 51–65. https://doi.org/10.1109/TITS.2014.2324414

Li, S., Wang, L., Yang, Z.Z., Ji, B., Qiao, F., Yang, Z.Z., Shiwu, L., Linhong, W., Zhifa, Y., Bingkui, J., Feiyan, Q., Zhongkai, Y., 2011. An active driver fatigue identification technique using multiple physiological features, in: Mechatronic Science, Electric Engineering and Computer (MEC), 2011 International Conference On. pp. 733–737. https://doi.org/10.1109/MEC.2011.6025569

Liang, Y., Lee, J.D., 2014. A hybrid Bayesian Network approach to detect driver cognitive distraction. Transp. Res. part C Emerg. Technol. 38, 146–155.

Liang, Y., Lee, J.D., 2010. Combining cognitive and visual distraction: Less than the sum of its parts. Accid. Anal. Prev. 42, 881–890. https://doi.org/10.1016/j.aap.2009.05.001

Liang, Y., Reyes, M.L., Lee, J.D., 2007. Real-time detection of driver cognitive distraction using support vector machines. IEEE Trans. Intell. Transp. Syst. 8, 340–350. https://doi.org/10.1109/TITS.2007.895298

Liu, P., Wang, Q., Gu, Y., 2009. Study on comparison of discretization methods. 2009

Int. Conf. Artif. Intell. Comput. Intell. AICI 2009 4, 380–384. https://doi.org/10.1109/AICI.2009.385

Lorenz, L., Kerschbaum, P., Schumann, J., 2014. Designing take over scenarios for automated driving How does augmented reality support the driver to get back into the loop?, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. pp. 1681–1685.

Loughborough University, 2018. Sport | About the University | Loughborough University [WWW Document]. URL https://www.lboro.ac.uk/about/sport/ (accessed 12.28.18).

Louppe, G., Wehenkel, L., Sutera, A., Geurts, P., 2013. Understanding variable importances in forests of randomized trees.

Louw, T., Madigan, R., Carsten, O., Merat, N., 2016. Were they in the loop during automated driving? Links between visual attention and crash potential. Inj. Prev. injuryprev--2016.

Louw, T.L., 2017. The Human Factors of Transitions in Highly Automated Driving.

Lu, Z., Winter, J.C.F. De, 2015. ScienceDirect A review and framework of control authority transitions in automated driving 00, 901–908.

Lyons, J.B., 2013. Being transparent about transparency, in: AAAI Spring Symposium.

Mabbott, N., 2003. ARRB Pro-active fatigue management system. acrs.org.au.

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations.

Mallis, M.M., Mejdal, S., Nguyen, T.T., Dinges, D.F., 2004. Summary of the key features of seven biomathematical models of human fatigue and performance. Aviat. Sp. Environ. Med. 75. https://doi.org/http://dx.doi.org/10.1016/j.biopsych.2010.03.008

Markkula, G., Benderius, O., Wolff, K., Wahde, M., 2012. A review of near-collision driver behavior models, in: Human Factors. pp. 1117–1143. https://doi.org/10.1177/0018720812448474

Marquaridt, D.W., 1970. Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation. Technometrics 12, 591–612. https://doi.org/10.1080/00401706.1970.10488699

Marquart, G., Cabrall, C., de Winter, J., 2015. Review of Eye-related Measures of Drivers' Mental Workload. Procedia Manuf. 3, 2854–2861. https://doi.org/10.1016/J.PROMFG.2015.07.783

Marulanda-Carter, L., Jackson, T.W., 2012. Effects of e-mail addiction and interruptions on employees. J. Syst. Inf. Technol. 14, 82–94. https://doi.org/10.1108/13287261211221146

Masala, G.L., Grosso, E., 2014. Real time detection of driver attention: Emerging solutions based on robust iconic classifiers and dictionary of poses. Transp. Res. part C Emerg. Technol. 49, 32–42.

May, J.G., Kennedy, R.S., Williams, M.C., Dunlap, W.P., Brannan, J.R., 1990. Eye movement indices of mental workload. Acta Psychol. (Amst). 75, 75–89. https://doi.org/10.1016/0001-6918(90)90067-P

Mehler, B., Reimer, B., Coughlin, J.F., Dusek, J.A., 2009. Impact of Incremental Increases in Cognitive Workload on Physiological Arousal and Performance in Young Adult Drivers. Transp. Res. Rec. J. Transp. Res. Board 2138, 6–12. https://doi.org/10.3141/2138-02

Melcher, V., Rauh, S., Diederichs, F., Bauer, W., 2015. ScienceDirect Take-Over Requests for automated driving 00, 4219–4225.

Melnicuk, V., Birrell, S.A., Konstantopoulos, P., Crundall, E., Jennings, P.A., 2016. JLR heart: employing wearable technology in non-intrusive driver state monitoring. Preliminary study.

Merat, N., de Waard, D., 2014. Human factors implications of vehicle automation: Current understanding and future directions. Transp. Res. part F traffic Psychol. Behav. 27, 193–195.

Merat, N., Jamson, A.H., Lai, F.C.H., Carsten, O., 2012. Highly automated driving, secondary task performance, and driver state. Hum. Factors J. Hum. Factors Ergon. Soc. 54, 762–771. https://doi.org/10.1177/0018720812442087

Merat, N., Jamson, A.H., Lai, F.C.H., Daly, M., Carsten, O.M.J., 2014. Transition to manual : Driver behaviour when resuming control from a highly automated vehicle. Transp. Res. Part F Psychol. Behav. 27, 274–282. https://doi.org/10.1016/j.trf.2014.09.005

Miyaji, M., Kawanaka, H., Oguri, K., 2010. Effect of pattern recognition features on detection for driver's cognitive distraction, in: 13th International IEEE Conference on Intelligent Transportation Systems. IEEE, pp. 605–610. https://doi.org/10.1109/ITSC.2010.5624966

Miyaji, M., Kawanaka, H., Oguri, K., 2009. Driver's cognitive distraction detection using physiological features by the adaboost, in: 2009 12th International IEEE Conference on Intelligent Transportation Systems. pp. 1–6. https://doi.org/10.1109/ITSC.2009.5309881

Mok, B., Johns, M., Miller, D., Ju, W., 2017. Tunneled In: Drivers with Active Secondary Tasks Need More Time to Transition from Automation. Proc. 2017 CHI Conf. Hum. Factors Comput. Syst. - CHI '17 2840–2844. https://doi.org/10.1145/3025453.3025713

Monk, C.A., Boehm-Davis, D.A., Mason, G., Trafton, J.G., 2004. Recovering from interruptions: Implications for driver distraction research. Hum. Factors J. Hum. Factors Ergon. Soc. 46, 650–663.

Morad, Y., Lemberg, H., Yofe, N., Dagan, Y., 2000. Pupillography as an objective indicator of fatigue. Curr. Eye Res. 21, 535–542.

Morgan, P., Alford, C., Parkhurst, G., 2016. Handover issues in autonomous driving: A literature review.

Nakayama, M., 2006. Influence of blink on pupillary indices, in: 2006 IEEE Biomedical Circuits and Systems Conference. pp. 29–32.

Nanopoulos, A., Alcock, B., Manolpoulos, Y., 2001. Feature-based classification of time-series data. researchgate.net. https://doi.org/10.1016/S0098-1354(02)00162-X

Naujoks, F., Purucker, C., Neukum, A., 2016. Secondary task engagement and vehicle automation – Comparing the effects of different automation levels in an on-road experiment. Transp. Res. Part F Traffic Psychol. Behav. 38, 67–82. https://doi.org/10.1016/J.TRF.2016.01.011

Neter, J., Kutner, M.H., Nachtsheim, C.J., Wasserman, W., 1996. Applied linear regression models. 1996. Irwin, Chicago, USA. p 1050.

Neubauer, C., Matthews, G., Langheim, L., Saxby, D., 2012. Fatigue and voluntary utilization of automation in simulated driving. Hum. Factors J. Hum. Factors Ergon. Soc. 54, 734–746.

Neumann, D.L., Lipp, O. V, 2002. Spontaneous and reflexive eye activity measures of mental workload. Aust. J. Psychol. 54, 174–179.

New York Times, 1925. HOUDINI SUBPOENAED WAITING TO BROADCAST; Magician Must Appear in Court on Charge That He Was Disorderly in Plaintiff's Office. - The New York Times [WWW Document]. URL https://www.nytimes.com/1925/07/23/archives/houdini-subpoenaed-waiting-to-broadcast-magician-must-appear-in.html (accessed 11.30.18).

Nguyen, T. Le, Gsponer, S., Ilie, I., Ifrim, G., 2018. Interpretable Time Series

Classification using All-Subsequence Learning and Symbolic Representations in Time and Frequency Domains.

NHS, 2015a. electrocardiography nhs. NHS.

NHS, 2015b. Electroencephalography – An Overview. NHS.

NHTSA, 2013. U.S. Department of Transportation Releases Policy on Automated Vehicle Development. U.S. Dep. Transp. Releases Policy Autom. Veh. Dev.

Nicodemus, K.K., Malley, J.D., 2009. Predictor correlation impacts machine learning algorithms: Implications for genomic studies. Bioinformatics 25, 1884–1890. https://doi.org/10.1093/bioinformatics/btp331

Nilsson, J., Falcone, P., Vinter, J., 2015. Safe Transitions From Automated to Manual Driving Using Driver Controllability Estimation. IEEE Trans. Intell. Transp. Syst. 16, 1806–1816.

O'Brien, R.M., 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. Qual. Quant. 41, 673–690. https://doi.org/10.1007/s11135-006-9018-6

Papageorgiou, M., Diakaki, C., Nikolos, I., Ntousakis, I., Papamichail, I., Roncoli, C., 2015. Freeway Traffic Management in Presence of Vehicle Automation and Communication Systems (VACS). Springer, Cham, pp. 205–214. https://doi.org/10.1007/978-3-319-19078-5_18

Parasuraman, R., Manzey, D.H., 2010. Complacency and bias in human use of automation: An attentional integration. Hum. Factors 52, 381–410. https://doi.org/10.1177/0018720810376055

Park, S., Trivedi, M., 2005. Driver activity analysis for intelligent vehicles: issues and development framework, in: IEEE Proceedings. Intelligent Vehicles Symposium, 2005. pp. 644–649.

Patrick Schäfer, 2017. patrickzib/SFA: Scalable Time Series Data Analytics [WWW Document]. URL https://github.com/patrickzib/SFA (accessed 11.26.18).

Patro, S.G.K., Sahu, K.K., 2015. Normalization: A Preprocessing Stage.

Paxion, J., Galy, E., Berthelon, C., 2014. Mental workload and driving. Front. Psychol. 5.

Payre, W., Cestac, J., Delhomme, P., 2014. Intention to use a fully automated car : Attitudes and a priori acceptability. Transp. Res. Part F Psychol. Behav. 27, 252–263. https://doi.org/10.1016/j.trf.2014.04.009

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Petermann-Stock, I., Hackenberg, L., Muhr, T., Mergl, C., 2013. Wie lange braucht der Fahrer--eine Analyse zu {Ü}bernahmezeiten aus verschiedenen Nebent{ä}tigkeiten w{ä}hrend einer hochautomatisierten Staufahrt. 6. Tagung Fahrerassistenzsysteme. Der Weg zum Autom. Fahr.

Pfleging, B., Fekety, D.K., Schmidt, A., Kun, A.L., 2016. A Model Relating Pupil Diameter to Mental Workload and Lighting Conditions, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16. ACM Press, New York, New York, USA, pp. 5776–5788. https://doi.org/10.1145/2858036.2858117

Plowman, S., Smith, D., 2007. Exercise Physiology for Health, Fitness, and Performance. Wolters Kluwer Health/Lippincott Williams & Wilkins. https://doi.org/10.1080/02705060.2008.9664239

Poczter, SL & Jankovic, L., 2014. "The Google Car: driving toward a better future?",. J. Bus. Case Stud. (Online), vol. 10, no. 1, p. 7.

Polar, 2016. Polar A360 product support | Polar UK [WWW Document]. URL https://support.polar.com/uk-en/support/polar_a360_product_support (accessed

9.27.18).

Quinlan, 1986. Induction of Decision Trees. Mach. Learn. 1, 81–106.

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., Bengler, K., 2014. How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. pp. 2063–2067.

Radlmayr, J., Ratter, M., Feldhütter, A., Körber, M., 2019. Take-overs in Level 3 automated driving – Proposal of the take-over performance score (TOPS) 821, 1–10. https://doi.org/10.1007/978-3-319-96080-7

Raileanu, L.E., Stoffel, K., 2004. Theoretical comparison between the Gini Index and Information Gain criteria *. Ann. Math. Artif. Intell. 41, 77–93.

Rakotonirainy, A., Schroeter, R., Soro, A., 2014. Three social car visions to improve driver behaviour. Pervasive Mob. Comput. 14, 147–160. https://doi.org/10.1016/j.pmcj.2014.06.004

Reason, J.T., 1990. Human error. Cambridge University Press.

Recarte, M.A., Nunes, L.M., 2003. Mental workload while driving: effects on visual search, discrimination, and decision making. J. Exp. Psychol. Appl. 9, 119.

Recarte, M.A., Nunes, L.M., 2000. Effects of verbal and spatial-imagery tasks on eye fixations while driving. J. Exp. Psychol. Appl. 6, 31–43. https://doi.org/10.1037/1076-898X.6.1.31

Regan, M.A., Hallett, C., Gordon, C.P., 2011. Driver distraction and driver inattention: Definition, relationship and taxonomy. Accid. Anal. Prev. 43, 1771–1781.

Reimer, B., Mehler, B., Coughlin, J.F., Godfrey, K.M., Tan, C., Bryan Reimer*, Bruce Mehler, Joseph F. Coughlin, K.M.G. and C.T., 2009. An on-road assessment of the impact of cognitive workload on physiological arousal in young adult drivers, in: Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications. pp. 115–118.

Reus, N. De, 1994. Assessment of Benefits and Drawbacks of Using Fuzzy Logic, Especially in Fire Control Systems. TNO Def. Res.

Reuschenbach, A., Wang, M., Ganjineh, T., Göhring, D., 2010. IDriver - Human machine interface for autonomous cars, in: Proceedings - 2011 8th International Conference on Information Technology: New Generations, ITNG 2011. pp. 435–440. https://doi.org/10.1109/ITNG.2011.83

ROSCOE, A.H., 1993. Heart rate as a psychophysiological measure for in-flight workload assessment. Ergonomics 36, 1055–1062. https://doi.org/10.1080/00140139308967977

Sabey, B.E., Taylor, H., 1980. The Known Risks We Run: The Highway, in: Societal Risk Assessment. Springer US, Boston, MA, pp. 43–70. https://doi.org/10.1007/978-1-4899-0445-4_3

SAE International, 2018. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. https://doi.org/10.4271/981198.Dingus

SAE International, 2016. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. https://doi.org/10.4271/981198.Dingus

SAE International, 2014. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. https://doi.org/10.2514/6.2006-3024

Saffarian, M., Winter, J.C.F. De, Happee, R., 2012. Automated Driving : Human-factors issues and design solutions 2296–2300.

Salmerón Gómez, R., García Pérez, J., López Martín, M.D.M., García, C.G., 2016.

Collinearity diagnostic applied in ridge estimation through the variance inflation factor. J. Appl. Stat. 43, 1831–1849. https://doi.org/10.1080/02664763.2015.1120712

Salthouse, T.A., 2009. When does age-related cognitive decline begin? Neurobiol. Aging 30, 507–514.

Salvucci, D.D., Bogunovich, P., 2010. Multitasking and monotasking, in: Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10. ACM Press, New York, New York, USA, p. 85. https://doi.org/10.1145/1753326.1753340

Santos, J., Merat, N., Mouta, S., Brookhuis, K., de Waard, D., 2005. The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. Transp. Res. Part F Traffic Psychol. Behav. 8, 135–146. https://doi.org/10.1016/J.TRF.2005.04.001

Sastry, K., Goldberg, D., Kendall, G., 2005. Genetic Algorithms, in: Search Methodologies. Springer US, Boston, MA, pp. 97–125. https://doi.org/10.1007/0-387-28356-0_4

Saxby, D.J., Matthews, G., Warm, J.S., Hitchcock, E.M., Neubauer, C., 2013. Active and passive fatigue in simulated driving: Discriminating styles of workload regulation and their safety impacts. J. Exp. Psychol. Appl. 19, 287–300. https://doi.org/10.1037/a0034386

Sayer, J.R., Devonshire, J.M., Flannagan, C.A., 2005. The effects of secondary tasks on naturalistic driving performance.

Schäfer, P., 2015. The BOSS is concerned with time series classification in the presence of noise. Data Min. Knowl. Discov. 29, 1505–1530. https://doi.org/10.1007/s10618-014-0377-7

Schäfer, P., Högqvist, M., 2012. SFA: a symbolic fourier approximation and index for similarity search in high dimensional datasets. Proc. 15th Int. Conf. Extending Database Technol. 516–527. https://doi.org/10.1145/2247596.2247656

Schäfer, P., Leser, U., 2017a. Multivariate Time Series Classification with WEASEL+MUSE. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Schäfer, P., Leser, U., 2017b. Fast and Accurate Time Series Classification with WEASEL. https://doi.org/10.1145/3132847.3132980

Schmitt, F., Korthauer, A., Manstetten, D., Bieg, H.-J., 2018. Predicting Strategies of Driving in Presence of Additional Visually Demanding Tasks: Inverse Optimal Control Estimation of Steering and Glance Behaviour Models, in: UR:BAN Human Factors in Traffic. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 183–204. https://doi.org/10.1007/978-3-658-15418-9_10

Schnittker, R., Marshall, S., Horberry, T., 2019. The co-design process of a decision support tool for airway management, in: Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018). Springer, Cham, pp. 111–120. https://doi.org/10.1007/978-3-319-96080-7

Schooler, J.W., Smallwood, J., Christoff, K., Handy, T.C., Reichle, E.D., Sayette, M.A., 2011. Meta-awareness, perceptual decoupling and the wandering mind. Trends Cogn. Sci. 15, 319–326. https://doi.org/10.1016/j.tics.2011.05.006

Sentinel, M., 1926. Phantom Auto'will tour city. Milwaukee Sentin. 4.

Seppäläinen, A.M.H., Landrigan, P.J., 1988. Neurophysiological approaches to the detection of early neurotoxicity in humans. CRC Crit. Rev. Toxicol. 18, 245–298.

Sigari, M.-H., Fathy, M., Soryani, M., 2013. A Driver Face Monitoring System for Fatigue and Distraction Detection. Int. J. Veh. Technol. 2013, 1–11. https://doi.org/10.1155/2013/263983

Sikander, G., Anwar, S., 2018. Driver Fatigue Detection Systems: A Review. IEEE Trans. Intell. Transp. Syst. 1–14. https://doi.org/10.1109/TITS.2018.2868499

Siu, K.-C., Chou, L.-S., Mayr, U., Van Donkelaar, P., Woollacott, M.H., 2008. Does inability to allocate attention contribute to balance constraints during gait in older adults? Journals Gerontol. Ser. A Biol. Sci. Med. Sci. 63, 1364–1369.

Smola, A., Vishwanathan, S.V.N.V.N., 2014. Introduction to machine learning. Methods Mol. Biol. 1107, 105–128. https://doi.org/10.1007/978-1-62703-748-8-7

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Solovey, E.T., Zec, M., Garcia Perez, E.A., Reimer, B., Mehler, B., 2014. Classifying driver workload using physiological and driving performance data, in: Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14. ACM Press, New York, New York, USA, pp. 4057–4066. https://doi.org/10.1145/2556288.2557068

Son, J., Park, M., Oh, H., 2012. Detecting Cognitive Workload Using Driving Performance and Eye Movement in a Driving Simulator. proceeding 11th Int. Symp. Adv. Veh. Control 1–4.

Srinivas, C., Reddy, B.R., Ramji, K., Naveen, R., 2014. Sensitivity Analysis to Determine the Parameters of Genetic Algorithm for Machine Layout. Procedia Mater. Sci. 6, 866–876. https://doi.org/10.1016/J.MSPRO.2014.07.104

Srinivas, M., Patnaik, L.M., 1994. Genetic algorithms: a survey. Computer (Long. Beach. Calif). 27, 17–26. https://doi.org/10.1109/2.294849

Stanton, N.A., Stewart, R., Harris, D., Houghton, R.J., Baber, C., McMaster, R., Salmon, P., Hoyle, G., Walker, G., Young, M.S., Linsell, M., Dymott, R., Green, D., 2006. Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. Ergonomics 49, 1288–1311. https://doi.org/10.1080/00140130600612762

Stern, J.A., Boyer, D., Schroeder, D., 1994. Blink rate: a possible measure of fatigue.

Hum. Factors J. Hum. Factors Ergon. Soc. 36, 285–297.

STISIM, 2018. Car driving simulator and software for occupational therapy, research, and training [WWW Document]. 2017. URL http://stisimdrive.com/ (accessed 12.2.18).

STISIM, 2007. SCENARIO DEFINITION LANGUAGE [WWW Document]. URL http://web.mit.edu/16.400/www/auto_sim/Help/SDL.htm (accessed 12.25.18).

Strayer, D.L., Drews, F.A., Crouch, D.J., 2006. A comparison of the cell phone driver and the drunk driver. Hum. factors J. Hum. factors Ergon. Soc. 48, 381–391.

Strayer, D.L., Drews, F.A., Johnson, W.A., 2003. Are We Being Driven to Distraction?

Summers, B.L., 2011. Software engineering reviews and audits. CRC Press.

Suykens, J.A.K., Vandewalle, J., 1999. Least Squares Support Vector Machine Classifiers. Neural Process. Lett. 9, 293–300. https://doi.org/10.1023/A:1018628609742

Tapia, E.M., Intille, S.S., Haskell, W., Larson, K.W.J., King, A., Friedman, R., 2007. Real-Time Recognition of Physical Activities and their Intensitiies Using Wireless Accelerometers and a Heart Monitor. Int. Symp. Wearable Comput. 37–40. https://doi.org/10.1109/ISWC.2007.4373774

Teh, E., Jamson, S., Carsten, O., Jamson, H., 2014. Temporal fluctuations in driving demand : The effect of traffic complexity on subjective measures of workload and driving performance. Transp. Res. Part F Psychol. Behav. 22, 207–217. https://doi.org/10.1016/j.trf.2013.12.005

Thomas, L.C., Gast, C., Grube, R., Craig, K., 2015. Fatigue Detection in Commercial Flight Operations: Results Using Physiological Measures. Procedia Manuf. 3, 2357–2364. https://doi.org/10.1016/j.promfg.2015.07.383

Tibshiranit, R., 1996. Regression Shrinkage and Selection via the Lasso. J. R. Stat.

Soc. Ser. B J. R. Stat. Soc. B 58, 267–288.

Tin Kam Ho, 1998. The random subspace method for constructing decision forests. IEEE Trans. Pattern Anal. Mach. Intell. 20, 832–844. https://doi.org/10.1109/34.709601

Tobii, 2016a. Tobii Pro Glasses 2.

Tobii, 2016b. Tobii Dynavox: We help you communicate – whatever it takes.

Tokuda, S., 2010. Using saccadic eye movements to measure mental workload. J. Vis. 8, 86–86. https://doi.org/10.1167/8.17.86

Urdan, T.C., 2005. Statistics in plain English 15.

van den Beukel, A.P., van der Voort, M.C., 2013. The influence of time-criticality on Situation Awareness when retrieving human control after automated driving, in: 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013). IEEE, pp. 2000–2005. https://doi.org/10.1109/ITSC.2013.6728523

Victor, T.W., 2005. Keeping Eye and Mind on the Road. Digit. Compr. Summ. Uppsala Diss. from Fac. Soc. Sci. 9 83 s.

Vogelpohl, T., Kühn, M., Hummel, T., Gehlert, T., Vollrath, M., 2018. Transitioning to manual driving requires additional time after automation deactivation. Transp. Res. Part F Traffic Psychol. Behav. 55, 464–482. https://doi.org/10.1016/J.TRF.2018.03.019

Volvo, 2013. Pilot Assist and Lane assistance | Volvo Cars UK [WWW Document]. URL https://www.volvocars.com/uk/support/article/3395a91b40bddd9ac0a801511916dab3 (accessed 11.28.18).

Walsham, G., 1995. The emergence of interpretivism in IS research. Inf. Syst. Res. 6, 376–394.

Walsorth, M.T., 1882. Twenty Questions: A Short Treatise on the Game to which are

Added a Code of Rules and Specimen Games for the Use of Beginners. Holt.

Wandtner, B., Schmidt, G., Schoemig, N., Kunde, W., 2018. Non-driving related tasks in highly automated driving – Effects of task modalities and cognitive workload on take-over performance Participants. Automot. meets Electron. Conf. 8–13.

Wang, R., Blackburn, G., Desai, M., Phelan, D., Gillinov, L., Houghtaling, P., Gillinov, M., 2016. Accuracy of Wrist-Worn Heart Rate Monitors. JAMA Cardiol.

Wang, Y., Fan, Y., Bhatt, P., Davatzikos, C., 2010. High-dimensional pattern regression using machine learning: From medical images to continuous clinical variables. Neuroimage 50, 1519–1535. https://doi.org/10.1016/J.NEUROIMAGE.2009.12.092

Wang, Z., Oates, T., 2014. Encoding Time Series as Images for Visual Inspection and Classification Using Tiled Convolutional Neural Networks.

Weber, M., 2014. Where to? A History of Autonomous Vehicles | Computer History Museum. Comput. Hist. Museum 13, 1–31.

Welch, D., Behrmann, E., 2018. Who's Winning the Self-Driving Car Race? - Bloomberg [WWW Document]. URL https://www.bloomberg.com/news/features/2018-05-07/who-s-winning-the-self-driving-car-race (accessed 11.28.18).

Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T., Vapnik, V., 2000. Feature selection for SVMs. Proc. 13th Int. Conf. Neural Inf. Process. Syst.

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.

Wijsman, J., Grundlehner, B., Liu, H., Hermens, H., Penders, J., 2011. Towards mental stress detection using wearable physiological sensors, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. IEEE, pp. 1798–1801. https://doi.org/10.1109/IEMBS.2011.6090512

Wille, M., Röwenstrunk, M., Debus, G., 2008. KONVOI: Electronically coupled truck convoys, in: Human Factors for Assistance and Automation. pp. 243–256.

Wilson, G.F., 2002. An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures. Int. J. Aviat. Psychol. 12, 3–18. https://doi.org/10.1207/S15327108IJAP1201_2

Wolcott, R., 2018. How Automation Will Change Work, Purpose, and Meaning [WWW Document]. URL https://hbr.org/2018/01/how-automation-will-change-work-purpose-and-meaning (accessed 11.27.18).

Wolpert, D.H., 2002. The Supervised Learning No-Free-Lunch Theorems, in: Soft Computing and Industry. Springer London, London, pp. 25–42. https://doi.org/10.1007/978-1-4471-0123-9_3

Wright, T.J., Samuel, S., Borowsky, A., Zilberstein, S., Fisher, D.L., 2016a. Experienced drivers are quicker to achieve situation awareness than inexperienced drivers in situations of transfer of control within a Level 3 autonomous environment, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. pp. 270–273.

Wright, T.J., Samuel, S., Borowsky, A., Zilberstein, S., Fisher, D.L., Swedish National, R., Transport Research, I., 2016b. Are experienced drivers quicker to regain full situation awareness in scenarios involving transfer of control from the automation to the driver? p. 3p.

Wu, J., Yao, L., Liu, B., 2018. An overview on feature-based classification algorithms for multivariate time series, in: 2018 3rd IEEE International Conference on

Cloud Computing and Big Data Analysis, ICCCBDA 2018. IEEE, pp. 32–38. https://doi.org/10.1109/ICCCBDA.2018.8386483

Yang, J., Honavar, V., 1998a. Feature Subset Selection Using a Genetic Algorithm, in: Feature Extraction, Construction and Selection. Springer US, Boston, MA, pp. 117–136. https://doi.org/10.1007/978-1-4615-5725-8_8

Yang, J., Honavar, V., 1998b. Feature subset selection using genetic algorithm. IEEE Intell. Syst. Their Appl. 13, 44–48. https://doi.org/10.1109/5254.671091

Yang, Y., Sun, H., Liu, T., Huang, G.-B., Sourina, O., 2015. Driver Workload Detection in On-Road Driving Environment Using Machine Learning. Springer, Cham, pp. 389–398. https://doi.org/10.1007/978-3-319-14066-7_37

Young, M.S., Stanton, N.A., 2002. Malleable Attentional Resources Theory: A New Explanation for the Effects of Mental Underload on Performance. Hum. Factors J. Hum. Factors Ergon. Soc. 44, 365–375. https://doi.org/10.1518/0018720024497709

Young, R.A., Hsieh, L., Seaman, S., 2013. The tactile detection response task: preliminary validation for measuring the attentional effects of cognitive load, in: Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design. pp. 71–77.

Yu, L., Sun, X., Zhang, K., 2011. Driving distraction analysis by ECG signals: an entropy analysis, in: International Conference on Internationalization, Design and Global Development. pp. 258–264. https://doi.org/10.1007/978-3-642-21660-2_29

Yu, L., Wang, S., Lai, K.K., 2006. An integrated data preparation scheme for neural network data analysis. IEEE Trans. Knowl. Data Eng. 18, 217–230.

Zeeb, K., Buchner, A., Schrauf, M., 2016. Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. Accid. Anal. Prev. 92, 230–239.

Zeeb, K., Buchner, A., Schrauf, M., 2015. What determines the take-over time? An integrated model approach of driver take-over after automated driving. Accid. Anal. Prev. 78, 212–221. https://doi.org/10.1016/j.aap.2015.02.023

Zhang, H., 2004. The Optimality of Naive Bayes, in: Florida Artificial Intelligence Research Society Conference. pp. 1–6. https://doi.org/10.1016/j.patrec.2005.12.001

Zhang, S., Chen, J., Lyu, F., Cheng, N., Shi, W., Shen, X., 2018. Vehicular Communication Networks in the Automated Driving Era. IEEE Commun. Mag. 56, 26–32. https://doi.org/10.1109/MCOM.2018.1701171

Zhang, Y., Hua, C., 2015. Driver fatigue recognition based on facial expression analysis using local binary patterns. Opt. - Int. J. Light Electron Opt. 126, 4501–4505. https://doi.org/10.1016/j.ijleo.2015.08.185

Zhang, Y., Owechko, Y., Zhang, J., 2004. Driver cognitive workload estimation: A data-driven perspective, in: Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference On. pp. 642–647.

Zhao, Y., Hryniewicki, M.K., Cheng, F., Fu, B., Zhu, X., 2019. Employee Turnover Prediction with Machine Learning: A Reliable Approach. Springer, Cham, pp. 737–758. https://doi.org/10.1007/978-3-030-01057-7_56

Zheng, A., Casari, A., 2018. Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists, in: O'Reilly. p. 218.

Zheng, Y., Hansen, J.H.L., 2016. Unsupervised driving performance assessment using free-positioned smartphones in vehicles, in: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). IEEE, pp. 1598–1603. https://doi.org/10.1109/ITSC.2016.7795771

Appendix A: Details of the Experiment Design

The experiment used the official scripting language of STISIM software. The following section include all the code used in the experiment.

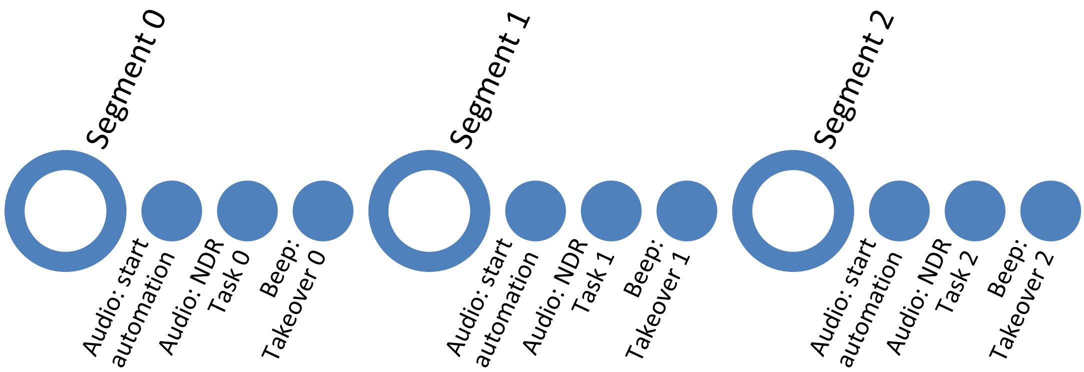


Figure 28: An illustration of the flow of the experiment.

The code is structured as illustrated in Figure 28. The main experiment is run through one of the six files named 'scenario_xx_yy_zz.EVT', where xx, yy, and zz are the names of the NDR tasks. Each file consists of three segments. For example, 'Scenario_email_attention_tqt.EVT' is a scenario file that requests drivers to check their emails at the first segment, pay attention to the road (control group) at the second segment and the TQT at the third segment.

The Segment is a structured code to demonstrates the movement of vehicles on the three lanes the ego vehicle is moving. There are three different segments that are designed to allow the ego vehicle to be on three different lanes. Those files are named 'segmentX' where X is 0, 1 or 2. Finally, the vehicles that flow on the opposite side of the road are coded on a file named 'opposite0.PDE'.

Each segment has three major events. The first event is an audio message to inform the driver that automation is starting. After few minutes, a message is played asking the driver to engage in a secondary task. This could be of the tasks: email, TQT or 'pay attention to the road'. After approximately 7 minutes, a beep is played to inform the drivers that the automated system is disabled, and the vehicle is asking the driver to takeover. This routine is coded in a file called 'takeover2.PDE'. Once the driver handles the vehicle to avoid the critical hazard, the driver enters the next segment repeating the same steps all over again.

The Takeover process starts with adding a vehicle that slows down the nearby lane to allow a space for the driver to perform a safe manoeuvre. Then, a beep sound is played to signal that the automated system is disabled.

The following tables include a copy of all code required to run the experiment on any STISIM-based driving simulator.

Scenario_email_attention_tqt.EVT

```
0, Previously Defined Events, segment0.PDE
60000, Previously Defined Events, segment1.PDE
120000, Previously Defined Events, segment2.PDE


-1 , -------------------------------- start automation at the beginning



4000, Play Recording, C:\STISIM3\Sound\move_far_right_lane.wav, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated_driving_starts.WAV, 0,
100
6500, Play Recording, C:\STISIM3\Sound\center_wheel.WAV, 0, 100

6500, LS, 50, 1000
6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, CV, 102, 2, -14{0}, 3


-1, -------------------------------- audio messages

100, Play Recording, C:\STISIM3\Sound\70_mph.wav, 0, 100
```

```
15000, Play Recording, C:\STISIM3\Sound\attention_road.wav, 0, 100

70000, Play Recording, C:\STISIM3\Sound\emails.wav, 0, 100

135000, Play Recording, C:\STISIM3\Sound\tqt_tasks.wav, 0, 100
```

Scenario_attention_tqt_email.EVT

```
0, Previously Defined Events, segment0.PDE
60000, Previously Defined Events, segment1.PDE
120000, Previously Defined Events, segment2.PDE


-1 , ------------------------------ start automation at the beginning



4000, Play Recording, C:\STISIM3\Sound\move_far_right_lane.wav, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated_driving_starts.WAV, 0,
100
6500, Play Recording, C:\STISIM3\Sound\center_wheel.WAV, 0, 100

6500, LS, 50, 1000
6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, CV, 102, 2, -14{0}, 3


-1, ------------------------------- audio messages

100, Play Recording, C:\STISIM3\Sound\70_mph.wav, 0, 100

15000, Play Recording, C:\STISIM3\Sound\attention_road.wav, 0, 100

70000, Play Recording, C:\STISIM3\Sound\tqt_tasks.wav, 0, 100

135000, Play Recording, C:\STISIM3\Sound\emails.wav, 0, 100
```

Scenario_email_attention_tqt.EVT

```
0, Previously Defined Events, segment0.PDE
60000, Previously Defined Events, segment1.PDE
120000, Previously Defined Events, segment2.PDE


-1 , ------------------------------ start automation at the beginning



4000, Play Recording, C:\STISIM3\Sound\move_far_right_lane.wav, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated_driving_starts.WAV, 0,
100
6500, Play Recording, C:\STISIM3\Sound\center_wheel.WAV, 0, 100

6500, LS, 50, 1000
6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, CV, 102, 2, -14{0}, 3


-1, ------------------------------- audio messages

100, Play Recording, C:\STISIM3\Sound\70_mph.wav, 0, 100

15000, Play Recording, C:\STISIM3\Sound\emails.wav, 0, 100

70000, Play Recording, C:\STISIM3\Sound\attention_road.wav, 0, 100

135000, Play Recording, C:\STISIM3\Sound\tqt_tasks.wav, 0, 100
```

Scenario_emails_tqt_attention.EVT

```
0, Previously Defined Events, segment0.PDE
60000, Previously Defined Events, segment1.PDE
120000, Previously Defined Events, segment2.PDE


-1 , ------------------------------ start automation at the beginning
```

```
4000, Play Recording, C:\STISIM3\Sound\move_far_right_lane.wav, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated_driving_starts.WAV, 0,
100
6500, Play Recording, C:\STISIM3\Sound\center_wheel.WAV, 0, 100

6500, LS, 50, 1000
6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, CV, 102, 2, -14{0}, 3


-1, -------------------------------- audio messages

100, Play Recording, C:\STISIM3\Sound\70_mph.wav, 0, 100

15000, Play Recording, C:\STISIM3\Sound\emails.wav, 0, 100

70000, Play Recording, C:\STISIM3\Sound\tqt_tasks.wav, 0, 100

135000, Play Recording, C:\STISIM3\Sound\attention_road.wav, 0, 100
```

Scenario_tqt_attention_emails.EVT

```
0, Previously Defined Events, segment0.PDE
60000, Previously Defined Events, segment1.PDE
120000, Previously Defined Events, segment2.PDE


-1 , ------------------------------- start automation at the beginning


4000, Play Recording, C:\STISIM3\Sound\move_far_right_lane.wav, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated_driving_starts.WAV, 0,
100
6500, Play Recording, C:\STISIM3\Sound\center_wheel.WAV, 0, 100

6500, LS, 50, 1000
6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, CV, 102, 2, -14{0}, 3
```

```
-1, -------------------------------- audio messages

100, Play Recording, C:\STISIM3\Sound\70_mph.wav, 0, 100

15000, Play Recording, C:\STISIM3\Sound\tqt_tasks.wav, 0, 100

70000, Play Recording, C:\STISIM3\Sound\attention_road.wav, 0, 100

135000, Play Recording, C:\STISIM3\Sound\emails.wav, 0, 100
```

Scenario_tqt_emails_attention.EVT

```
0, Previously Defined Events, segment0.PDE
60000, Previously Defined Events, segment1.PDE
120000, Previously Defined Events, segment2.PDE

-1 , ------------------------------- start automation at the beginning

4000, Play Recording, C:\STISIM3\Sound\move_far_right_lane.wav, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated_driving_starts.WAV, 0,
100
6500, Play Recording, C:\STISIM3\Sound\center_wheel.WAV, 0, 100

6500, LS, 50, 1000
6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, CV, 102, 2, -14{0}, 3

-1, -------------------------------- audio messages

100, Play Recording, C:\STISIM3\Sound\70_mph.wav, 0, 100

15000, Play Recording, C:\STISIM3\Sound\tqt_tasks.wav, 0, 100

70000, Play Recording, C:\STISIM3\Sound\emails.wav, 0, 100

135000, Play Recording, C:\STISIM3\Sound\attention_road.wav, 0, 100
```

Takeover2.PDE

```
-1, ----  lat where static onject should be at , lat where vehicle should start
automation, sound instruction for moving to lane, sound instruction of
automation start, lat where vehicle is located at


100, Playback Marker
500, Vehicles, 1000, @5, 90{0}, 1, F41, 1, 1, &
1{15}, 2, 10{0], , , 6, &
3000{20}, 2, -5{3}, , , 6

700, Vehicles, 1000, @5, 90{0}, 1, F44, 1, 1, &
1{15}, 2, 10{0], , , 6, &
2600{20}, 2, -2{3}, , , 3, &
3000{20}, 2, 150{0], , , 5


3000, Play Recording, C:\STISIM3\Sound\beep.WAV, 2, 100, 7, 0
-1, control vehicle
3000, CV, , 0, @1
3000, Begin Block Save, 1, 0.99, TakeOver, 1, 4, 5, 10, 11, 19, 23, 27, 28, 35, 36, 39

3000, Vehicles, 700, @1, 0{0}, 1, T7, 1, 1, &
3800{20}, 9, 0.5
-1, 3000, Static Object,   700, @1, 0, 90, 0, 0, C:\STISIM3\Data\Vehicles\Trucks
and SUVs\Chevy_Tracker_Purple.Mka
-1, 3000, Collision Block, 700, @1, 30, 10,, 100

-1, 4000, Play Recording, @3, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated_driving_starts.wav, 0, 100

6500, LS, 50, 1000

6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, End Block Save
6500, CV, 102, 2, @2, 3
6600, Play Recording, C:\STISIM3\Sound\center_wheel.wav , 0, 100
```

Scenario0.EVT

```
0, Previously Defined Events, segment0.PDE

-1, 0, Control Vehicle, 102, 2, -36{0}, 2
-1, 100, Previously Defined Events, takeover2.PDE, -36{0}, -14{0},
C:\STISIM3\Sound\Move rt lane.wav, C:\STISIM3\Sound\automated.WAV, -
26{0}



-1 , ------------------------------ start automation at the beginning



4000, Play Recording, C:\STISIM3\Sound\Move rt lane.wav, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated.WAV, 0, 100

6500, LS, 50, 1000
6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, CV, 102, 2, -14{0}, 3


0, Previously Defined Events, opposite.PDE
```

Scenario1.EVT

```
0, Previously Defined Events, segment1.PDE

-1, 0, Control Vehicle, 102, 2, -36{0}, 2
-1, 100, Previously Defined Events, takeover2.PDE, -36{0}, -14{0},
C:\STISIM3\Sound\Move rt lane.wav, C:\STISIM3\Sound\automated.WAV, -
26{0}



-1 , ------------------------------ start automation at the beginning



4000, Play Recording, C:\STISIM3\Sound\Move rt lane.wav, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated.WAV, 0, 100

6500, LS, 50, 1000
6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
```

4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, CV, 102, 2, -26{0}, 3

Scenario2.EVT

0, Previously Defined Events, segment2.PDE

-1, 0, Control Vehicle, 102, 2, -36{0}, 2
-1, 100, Previously Defined Events, takeover2.PDE, -36{0}, -14{0}, C:\STISIM3\Sound\Move rt lane.wav, C:\STISIM3\Sound\automated.WAV, -26{0}

-1 , ------------------------------ start automation at the beginning

4000, Play Recording, C:\STISIM3\Sound\Move rt lane.wav, 0, 100

5500, Play Recording, C:\STISIM3\Sound\automated.WAV, 0, 100

6500, LS, 50, 1000
6500, SIGN, 100, 1000, c:\arriva\arrivasigns\Speed_50.3ds, 1, 1, 0
4500, SOBJ, 2000, -
60,0,90,0,0,C:\ARRIVA\Arrivasigns\gantry\NSL\Gantry_NSL_GIF.mesh

6500, CV, 102, 2, -36{0}, 3

Segement0.PDE

-
takeovers***
-1, 0, Control Vehicle, 102, 2, -36{0}, 2
-1, 100, Previously Defined Events, takeover2.PDE, -36{0}, -14{0}, C:\STISIM3\Sound\Move rt lane.wav, C:\STISIM3\Sound\automated_driving_starts.WAV, -26{0}

51000, Previously Defined Events, takeover2.PDE, -14{0}, -26{0}, C:\STISIM3\Sound\Move rt lane.wav, C:\STISIM3\Sound\automated_driving_starts.WAV, -26{0}

```
-1, ---- lat where static object should be at , lat where vehicle should start
automation, sound instruction for moving to lane, sound instruction of
automation start, lat where stopping-vehicle is located at



-straight
sections*********************************************************************
***********************************
-1, 0, Roadway, 12, 6, 5, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 50,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
-100, Roadway, 12, 6, 4, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
-1, 0, Roadway, 12, 6, 4, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
-1, 200, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
0, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
```

```
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
12000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
24000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
36000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
42000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
```

-Curved sections***

1300, Curve, 1200, 1200, 1000, 300, -0.0008
9000, Curve, 1200, 600, 1000, 300, 0.0008
15000, Curve, 1200, 600, 1000, 300, 0.0008
20000, Curve, 1200, 1200, 1000, 300, 0.0008
29000, Curve, 1200, 600, 1000, 300, -0.0008
35000, Curve, 1200, 1200, 1000, 300, 0.0008
45000, Curve, 1200, 1200, 1000, 300, 0.0008
50000, Curve, 1200, 600, 1000, 300, -0.0008

-hill sections up and back down

-2000, Hill, 150, 0.02
3000, Hill, 150, -0.02
4000, Hill, 150, -0.02
5000, Hill, 150, 0.02
14000, Hill, 150, 0.02
16000, Hill, 150, -0.02
18000, Hill, 150, -0.02
22000, Hill, 150, 0.02
24000, Hill, 150, 0.02
28000, Hill, 150, -0.02
32000, Hill, 150, -0.02
37000, Hill, 150, 0.02
42000, Hill, 150, 0.02

-Embankments

-Bridges***

6000, Previously Defined Events, C:\ARRIVA\ARRIVATRAINING\sessions\Motorwaybridge_start_conversion.pde

-overhead gantrys**

110, Speed Limit, 80, 50000

```
-
cars*******************************************************************************************************************

-1 cars in same lane as driver


-3lane motorway p1 = speed with * meaning your speed p2= distance away
vechile appears p3=distance from median

-1,  -----------------black caddilac vehicle in front of the ego lane-----
0, Vehicles, 600, -14{0}, 0{3}, 1, T7, 0, 1, &
54000{15}, 2, 120{0], , , 10


1500, Vehicles, 1200, -26{0}, 115, 1, T18, 1, 1
1800, Vehicles, 1200, -26{0}, 0{3}, 1, S9, 1, 18
1900, Vehicles, 1200, -36{0}, 102{0}, 1, C15, 1, 1

2400, Vehicles, 1200, -36{0}, 102{0}, 1, S46, 0, 1

4000, Vehicles, 1200, -36{0}, 102{0}, 1, F40, 0, 1
4500, Vehicles, 1200, -26{0}, 105, 1, F39, 1, 1

8500, Vehicles, 1200, -26{0}, 112, 1, F39, 1, 1

-1, 9000, Vehicles, 1200, -36{0}, 102{0}, 1, F4, 0, 1
9500, Vehicles, 1200, -26{0}, 117, 1, T35, 1, 1

-10200, Vehicles, 1200, -14{0}, 10{3}, 1, T15, 1, 1
10900, Vehicles, 1200, -36{0}, 102{0}, 1, C6, 1, 1

11400, Vehicles, 1200, -36{0}, 85{0}, 1, S46, 0, 1
11500, Vehicles, 1200, -26{0}, 112{0}, 1, F39, 1, 1

-12000, Vehicles, 1200, -14{0}, 10{3}, 1, T2, 0, 1
12400, Vehicles, 1200, -36{0}, 90{0}, 1, S8, 0, 1

-13000, Vehicles, 1200, -14{0}, 10{3}, 1, T22, 0, 1
13800, Vehicles, 1200, -26{0}, 0{3}, 1, S39, 1, 1

14000, Vehicles, 1200, -36{0}, 90{0}, 1, S14, 0, 1

18300, Vehicles, 1200, -26{0}, 0{3}, 1, S23, 1, 1, &
57000{15}, 2, 50{3], , , 10

19200, Vehicles, 1200, -36{0}, 90{0}, 1, S6, 1, 1

20700, Vehicles, 1200, -36{0}, 102{0}, 1, F21, 1, 1
20800, Vehicles, 1200, -26{0}, 0{3}, 1, S16, 1, 1
```

```
21200, Vehicles, 1200, -36{0}, 102{0}, 1, C15, 1, 1

22200, Vehicles, 1200, -36{0}, 98{0}, 1, F30, 0, 1
22500, Vehicles, 1200, -26{0}, 112{0}, 1, T16, 1, 1, &
54000{15}, 2, 30{3], , , 10


-23200, Vehicles, 1200, -14{0}, 10{3}, 1, T15, 1, 1
23300, Vehicles, 1200, -36{0}, 95{0}, 1, S6, 1, 1
23800, Vehicles, 1200, -26{0}, 0{3}, 1, S4, 1, 1, &
54000{15}, 2, 30{3], , , 10

24500, Vehicles, 1200, -26{0}, 15{3}, 1, T7, 1, 1, &
54000{15}, 2, 30{3], , , 10

24700, Vehicles, 1200, -36{0}, 93{0}, 1, S6, 1, 1
28000{15}, 2, -10{3], , , 10, &
35000{15}, 2, 10{3], , , 10

25800, Vehicles, 1200, -26{0}, 14{3}, 1, F26, 1, 1, &
54000{15}, 2, 30{3], , , 10

25850, Vehicles, 1200, -36{0}, 90{3}, 1, C15, 1, 1, &
30000{15}, 2, -10{3], , , 10, &
40000{15}, 2, 10{3], , , 10, &
54000{15}, 2, 120, , , 10

26400, Vehicles, 1200, -36{0}, 90{0}, 1, S46, 0, 1
-1, 26500, Vehicles, 1200, -26{0}, 102{0}, 1, F3, 1, 1, &
54000{15}, 2, 30{3], , , 10

27000, Vehicles, 1200, -36{0}, 98{0}, 1, F7, 0, 1
27800, Vehicles, 1200, -26{0}, 0{3}, 1, F4, 1, 1, &
32000{15}, 2, -10{3], , , 4, &
40000{15}, 2, 120{0], , , 10

30500, Vehicles, 1200, -26{0}, 1{3}, 1, S51, 1, 1, &
54000{15}, 2, 10{0], , , 10


31600, Vehicles, 1200, -14{0}, 10{3}, 1, T15, 1, 1
31900, Vehicles, 1200, -36{0}, 96{3}, 1, S23, 1, 1
31700, Vehicles, 1200, -26{0}, -10{3}, 1, S10, 1, 1, &
32000{15}, 2, 10{3], , , 10, &
45000{15}, 2, 120{0], , , 10

32400, Vehicles, 1200, -36{0}, 95{0}, 1, S46, 0, 1
32500, Vehicles, 1200, -26{0}, 112{0}, 1, S53, 1, 1, &
54000{15}, 2, 120{0], , , 6
32800, Vehicles, 1200, -26{0}, 0{3}, 1, F30, 1, 1, &
54000{15}, 2, 120{0], , , 10
```

```
32900, Vehicles, 1200, -36{0}, 90{0}, 1, S23, 1, 1

33100, Vehicles, 1500, 26{0}, 112{0}, 1, T15, 1, 1, &
54000{15}, 2, 120{0], , , 10
44450, Vehicles, 1500, 36{0}, 112{0}, 1, E9, 1, 1, &
54000{15}, 2, 120{0], , , 10
44978, Vehicles, 1500, 26{0}, 112{0}, 1, F56, 1, 1, &
54000{15}, 2, 120{0], , , 10

33600, Vehicles, 1200, -14{0}, 10{3}, 1, T15, 1, 1
33900, Vehicles, 1200, -36{0}, 0{3}, 1, S9, 1, 1


34500, Vehicles, 1200, -36{0}, 95{0}, 1, T26, 1, 1, &
54000{15}, 2, 120{0], , , 10

35800, Vehicles, 1200, -36{0}, 90{0}, 1, F5, 1, 1, &
54000{15}, 2, 120{0], , , 10

36800, Vehicles, 1200, -36{0}, 95{0}, 1, T23, 1, 1, &
54000{15}, 2, 120{0], , , 10

37800, Vehicles, 1200, -36{0}, 98{0}, 1, F19, 1, 1, &
54000{15}, 2, 120{0], , , 10

38800, Vehicles, 1200, -36{0}, 90{0}, 1, S48, 1, 1, &
54000{15}, 2, 120{0], , , 10

40100, Vehicles, 1200, -36{0}, 100{0}, 1, S53, 1, 1, &
54000{15}, 2, 120{0], , , 10
40400, Vehicles, 1200, -26{0}, 95{0}, 1, S5, 1, 1, &
54000{15}, 2, 60{0], , , 10


41100, Vehicles, 1200, -36{0}, 95{0}, 1, F43, 1, 1, &
54000{15}, 2, 120{0], , , 10
42800, Vehicles, 1200, -36{0}, 95{0}, 1, F26, 1, 1, &
54000{15}, 2, 120{0], , , 10

43100, Vehicles, 1200, -36{0}, 95{0}, 1, F22, 1, 1, &
54000{15}, 2, 120{0], , , 10
43400, Vehicles, 1200, -36{0}, 95{0}, 1, F4, 1, 1, &
56000{15}, 2, -20{3], , , 20


-1 cars on opposite side of road, -------------------------------------------------------------
0, Previously Defined Events, opposite.PDE


-1 crash
barriers*********************************************************************
*************************************
```

```
0, Barriers, 0, 0, 0{0}, 0{0}, 10000, 10
50000, Barriers, 0, 0, 0{0}, 0{0}, 6000, 10

-1,-----------------------------------------------------------------------------------------------------


-1 middle left facing

0, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
3000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
6000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
9000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
12000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
15000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
18000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
21000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
24000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
27000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
30000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
33000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
36000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
39000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
42000, Static Object, 0, -0.1{0}, 0, 0, *, *,
```

```
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
45000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
48000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
51000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10




0, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
1000, Tree, 55, 0, *1~3, 80{0}, 100{0}, 2
1000, Tree, 55, 0, *1~3, -80{0}, -100{0}, 2
```

Segment1.PDE

```
-
takeovers******************************************************************
********************************************
-1, 0, Control Vehicle, 102, 2, -36{0}, 2
-1, 100, Previously Defined Events, takeover2.PDE, -36{0}, -14{0},
C:\STISIM3\Sound\Move rt lane.wav,
C:\STISIM3\Sound\automated_driving_starts.WAV, -26{0}

51000, Previously Defined Events, takeover2.PDE, -26{0}, -36{0},
C:\STISIM3\Sound\Move rt lane.wav,
C:\STISIM3\Sound\automated_driving_starts.WAV, -36{0}

-1, ----  lat where static object should be at , lat where vehicle should start
automation, sound instruction for moving to lane, sound instruction of
automation start, lat where stopping-vehicle is located at



-straight
sections*********************************************************************
**********************************
-1, 0, Roadway, 12, 6, 5, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
```

```
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 50,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
-100, Roadway, 12, 6, 4, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
-1, 0, Roadway, 12, 6, 4, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
-1, 200, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
0, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
12000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
```

```
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
24000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
36000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
42000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0




-Curved
sections***********************************************************************
*************************************
1300, Curve, 1200, 1200, 1000, 300, -0.0008
9000, Curve, 1200, 600, 1000, 300, 0.0008
15000, Curve, 1200, 600, 1000, 300, 0.0008
20000, Curve, 1200, 1200, 1000, 300, 0.0008
29000, Curve, 1200, 600, 1000, 300, -0.0008
35000, Curve, 1200, 1200, 1000, 300, 0.0008
45000, Curve, 1200, 1200, 1000, 300, 0.0008
50000, Curve, 1200, 600, 1000, 300, -0.0008
```

-hill sections up and back down

-2000, Hill, 150, 0.02
3000, Hill, 150, -0.02
4000, Hill, 150, -0.02
5000, Hill, 150, 0.02
14000, Hill, 150, 0.02
16000, Hill, 150, -0.02
18000, Hill, 150, -0.02
22000, Hill, 150, 0.02
24000, Hill, 150, 0.02
28000, Hill, 150, -0.02
32000, Hill, 150, -0.02
37000, Hill, 150, 0.02
42000, Hill, 150, 0.02

-Embankments

-
Bridges***

6000, Previously Defined Events,
C:\ARRIVA\ARRIVATRAININGs\sessions\Motorwaybridge_start_conversion.pd
e

-overhead
gantrys**

110, Speed Limit, 80, 50000

-
cars***

-1 cars in same lane as driver

-3lane motorway p1 = speed with * meaning your speed p2= distance away vechile appears p3=distance from median

```
-1,  ----------------black caddilac vehicle in front of the ego lane-----
0, Vehicles, 600, -26{0}, 0{3}, 1, T7, 0, 1, &
54000{15}, 2, 120{0], , , 10


1500, Vehicles, 1200, -14{0}, 115, 1, T18, 1, 1

1800, Vehicles, 1200, -14{0}, 15{3}, 1, S9, 1, 18

1900, Vehicles, 1200, -14{0}, 20{3}, 1, C15, 1, 1


2400, Vehicles, 1200, -36{0}, 102{0}, 1, S46, 0, 1

4000, Vehicles, 1200, -36{0}, 102{0}, 1, F40, 0, 1
4500, Vehicles, 1200, -26{0}, 105, 1, F39, 1, 1

8500, Vehicles, -600, -14{0}, 130, 1, F39, 1, 1

-1, 9000, Vehicles, 1200, -36{0}, 102{0}, 1, F4, 0, 1
9500, Vehicles, -600, -14{0}, 120, 1, T35, 1, 1

-10200, Vehicles, 1200, -26{0}, 10{3}, 1, T15, 1, 1
-1, 10900, Vehicles, 1200, -36{0}, 102{0}, 1, C6, 1, 1

11500, Vehicles, -1200, -14{0}, 112{0}, 1, F39, 1, 1

-12000, Vehicles, 1200, -26{0}, 10{3}, 1, T2, 0, 1
12400, Vehicles, 1200, -36{0}, 90{0}, 1, S8, 0, 1

-13000, Vehicles, 1200, -26{0}, 10{3}, 1, T22, 0, 1
13800, Vehicles, -600, -14{0}, 20{3}, 1, S39, 1, 1

14000, Vehicles, 1200, -36{0}, 90{0}, 1, S14, 0, 1

18300, Vehicles, -600, -14{0}, 20{3}, 1, S23, 1, 1, &
57000{15}, 2, 50{3], , , 10

19200, Vehicles, 1200, -36{0}, 90{0}, 1, S6, 1, 1

20800, Vehicles, -600, -14{0}, 20{3}, 1, S16, 1, 1

-1, 21200, Vehicles, 1200, -36{0}, 102{0}, 1, C15, 1, 1

22200, Vehicles, 1200, -36{0}, 98{0}, 1, F30, 0, 1
22500, Vehicles, -600, -14{0}, 120{0}, 1, T16, 1, 1, &
54000{15}, 2, 30{3], , , 10


-23200, Vehicles, 1200, -26{0}, 10{3}, 1, T15, 1, 1
23300, Vehicles, 1200, -36{0}, 95{0}, 1, S6, 1, 1
23800, Vehicles, -600, -14{0}, 120{0}, 1, S4, 1, 1, &
```

```
54000{15}, 2, 30{3], , , 10

24500, Vehicles, -600, -14{0}, 20{3}, 1, T7, 1, 1, &
54000{15}, 2, 30{3], , , 10

24700, Vehicles, 1200, -36{0}, 93{0}, 1, S6, 1, 1
28000{15}, 2, -10{3], , , 10, &
35000{15}, 2, 10{3], , , 10

25800, Vehicles, -600, -14{0}, 20{3}, 1, F26, 1, 1, &
54000{15}, 2, 30{3], , , 10

-1, 25850, Vehicles, 1200, -36{0}, 90{3}, 1, C15, 1, 1, &
30000{15}, 2, -10{3], , , 10, &
40000{15}, 2, 10{3], , , 10, &
54000{15}, 2, 120, , , 10

26400, Vehicles, 1200, -36{0}, 90{0}, 1, S46, 0, 1
-1, 26500, Vehicles, -600, -14{0}, 120{0}, 1, F3, 1, 1, &
54000{15}, 2, 30{3], , , 10

27000, Vehicles, 1200, -36{0}, 98{0}, 1, F7, 0, 1
27800, Vehicles, -600, -14{0}, 20{3}, 1, F4, 1, 1

30500, Vehicles, -600, -14{0}, 20{3}, 1, S51, 1, 1


-1, 31600, Vehicles, 1200, -26{0}, 10{3}, 1, T15, 1, 1
31900, Vehicles, 1200, -36{0}, 96{3}, 1, S23, 1, 1
31700, Vehicles, -600, -14{0}, 20{3}, 1, S10, 1, 1, &
32000{15}, 2, 10{3], , , 10, &
45000{15}, 2, 160{0], , , 10

32400, Vehicles, 1200, -36{0}, 95{0}, 1, S46, 0, 1
32500, Vehicles, -600, -14{0}, 120{0}, 1, S53, 1, 1, &
54000{15}, 2, 120{0], , , 6
32800, Vehicles, -600, -14{0}, 120{3}, 1, F30, 1, 1, &
54000{15}, 2, 120{0], , , 10
32900, Vehicles, 1200, -36{0}, 90{0}, 1, S23, 1, 1

33900, Vehicles, 1200, -36{0}, 0{3}, 1, S9, 1, 1, &
54000{15}, 2, 150{0], , , 10

34500, Vehicles, 1200, -36{0}, 95{0}, 1, T26, 1, 1, &
54000{15}, 2, 150{0], , , 10

35800, Vehicles, 1200, -36{0}, 90{0}, 1, F5, 1, 1, &
54000{15}, 2, 150{0], , , 10

36800, Vehicles, 1200, -36{0}, 95{0}, 1, T23, 1, 1, &
54000{15}, 2, 150{0], , , 10
```

```
37800, Vehicles, 1200, -36{0}, 98{0}, 1, F19, 1, 1, &
54000{15}, 2, 150{0], , , 10

38800, Vehicles, 1200, -36{0}, 90{0}, 1, S48, 1, 1

40100, Vehicles, 1200, -36{0}, 100{0}, 1, S53, 1, 1, &
54000{15}, 2, 60{0], , , 10
40400, Vehicles, -600, -14{0}, 120{0}, 1, S5, 1, 1


41100, Vehicles, 1200, -36{0}, 95{0}, 1, F43, 1, 1, &
54000{15}, 2, 60{0], , , 10
42800, Vehicles, 1200, -36{0}, 95{0}, 1, F26, 1, 1, &
54000{15}, 2, 60{0], , , 10

43100, Vehicles, 1200, -36{0}, 95{0}, 1, F22, 1, 1, &
54000{15}, 2, 60{0], , , 10
43400, Vehicles, 1200, -36{0}, 95{0}, 1, F4, 1, 1, &
56000{15}, 2, -20{3], , , 20


-1 cars on opposite side of road, ----------------------------------------------------------
-1, 0, Previously Defined Events, opposite.PDE


-1 crash
barriers***************************************************************************
***************************************

-1,-----------------------------------------------------------------------------------------------


-1 middle left facing

0, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
3000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
6000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
9000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
12000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
15000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
```

```
18000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
21000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
24000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
27000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
30000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
33000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
36000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
39000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
42000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
45000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
48000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
51000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10




0, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
1000, Tree, 55, 0, *1~3, 80{0}, 100{0}, 2
1000, Tree, 55, 0, *1~3, -80{0}, -100{0}, 2
```

Segment2.PDE

```
-
takeovers*****************************************************************
*******************************************

51000, Previously Defined Events, takeover2.PDE, -36{0}, -26{0},
C:\STISIM3\Sound\Move rt lane.wav,
C:\STISIM3\Sound\automated_driving_starts.WAV, -26{0}

-1, ---- lat where static object should be at , lat where vehicle should start
automation, sound instruction for moving to lane, sound instruction of
automation start, lat where stopping-vehicle is located at



-straight
sections*****************************************************************
*********************************
-1, 0, Roadway, 12, 6, 5, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 50,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
-100, Roadway, 12, 6, 4, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
-1, 0, Roadway, 12, 6, 4, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\dirt02.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
-1, 200, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
```

```
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
0, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
12000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
24000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
36000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0
42000, Roadway, 12, 6, 3, 1, 0.7, 6.5, 13, 0.33, 0.33, 100, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 12,
```

C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 4,
C:\STISIM3\Data\Textures\Road03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0, 0, 100,
C:\STISIM3\Data\Textures\Grass03.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0,
C:\STISIM3\Data\Textures\grass08.Jpg, 255/255/255, 12, 0, 12,
C:\arriva\gif_files\SteveMedianl.gif, , 10, 0

-Curved
sections**

1300, Curve, 1200, 1200, 1000, 300, -0.0008
9000, Curve, 1200, 600, 1000, 300, 0.0008
15000, Curve, 1200, 600, 1000, 300, 0.0008
20000, Curve, 1200, 1200, 1000, 300, 0.0008
29000, Curve, 1200, 600, 1000, 300, -0.0008
35000, Curve, 1200, 1200, 1000, 300, 0.0008
45000, Curve, 1200, 1200, 1000, 300, 0.0008
50000, Curve, 1200, 600, 1000, 300, -0.0008

-hill sections up and back down

-2000, Hill, 150, 0.02
3000, Hill, 150, -0.02
4000, Hill, 150, -0.02
5000, Hill, 150, 0.02
14000, Hill, 150, 0.02
16000, Hill, 150, -0.02
18000, Hill, 150, -0.02
22000, Hill, 150, 0.02
24000, Hill, 150, 0.02
28000, Hill, 150, -0.02
32000, Hill, 150, -0.02
37000, Hill, 150, 0.02
42000, Hill, 150, 0.02

-Embankments

-
Bridges**

6000, Previously Defined Events,

```
C:\ARRIVA\ARRIVATRAINING\sessions\Motorwaybridge_start_conversion.pd
e




-overhead
gantrys**********************************************************************
*******************************
110, Speed Limit, 70, 50000


-
cars**********************************************************************
***********************************

-1 cars in same lane as driver


-3lane motorway p1 = speed with * meaning your speed p2= distance away
vechile appears p3=distance from median


-1,  -----------------black caddilac vehicle in front of the ego lane-----
0, Vehicles, 600, -36{0}, 0{3}, 1, T7, 0, 1, &
54000{15}, 2, 120{0], , , 10


1000, Vehicles, 1200, -14{0}, 15{3}, 1, S7, 1, 1
1200, Vehicles, 1200, -14{0}, 15{3}, 1, F37, 1, 1
1300, Vehicles, 1200, -14{0}, 15{3}, 1, S16, 1, 1
1400, Vehicles, -600, -14{0}, 25{3}, 1, S9, 1, 1


1500, Vehicles, 1200, -26{0}, -5{3}, 1, T18, 1, 1
1500, Vehicles, -600, -14{0}, 20{3}, 1, S46, 1, 1

1600, Vehicles, -600, -14{0}, 20{3}, 1, T35, 1, 1

1800, Vehicles, 1200, -26{0}, -5{3}, 1, S9, 1, 18
1800, Vehicles, 1200, -14{0}, 15{3}, 1, S9, 1, 18

1900, Vehicles, 1200, -14{0}, 20{3}, 1, C15, 1, 1


4500, Vehicles, 1200, -26{0}, -5{3}, 1, F39, 1, 1

8500, Vehicles, 1200, -26{0}, -2{3}, 1, F39, 1, 1
8500, Vehicles, -600, -14{0}, 130, 1, F37, 1, 1
```

```
9500, Vehicles, 1200, -26{0}, -1{3}, 1, T35, 1, 1
9500, Vehicles, -600, -14{0}, 120, 1, S8, 1, 1

-10200, Vehicles, 1200, -14{0}, 15{3}, 1, T15, 1, 1

11500, Vehicles, 1200, -26{0}, 105{0}, 1, F39, 1, 1
11500, Vehicles, -1200, -14{0}, 120{0}, 1, S9, 1, 1



13800, Vehicles, 1200, -26{0}, 0{3}, 1, S39, 1, 1
13800, Vehicles, -600, -14{0}, 20{3}, 1, S6, 1, 1


18300, Vehicles, 1200, -26{0}, 0{3}, 1, S23, 1, 1, &
57000{15}, 2, 50{3], , , 10
18300, Vehicles, -600, -14{0}, 20{3}, 1, T35, 1, 1

20800, Vehicles, 1200, -26{0}, 0{3}, 1, S16, 1, 1
20800, Vehicles, -600, -14{0}, 20{3}, 1, T7, 1, 1

22500, Vehicles, 1200, -26{0}, 112{0}, 1, T16, 1, 1, &
54000{15}, 2, 30{3], , , 10
22500, Vehicles, -600, -14{0}, 120{0}, 1, S6, 1, 1


23800, Vehicles, 1200, -26{0}, 0{3}, 1, S4, 1, 1, &
54000{15}, 2, 30{3], , , 10
23800, Vehicles, -600, -14{0}, 120{0}, 1, F30, 1, 1

24500, Vehicles, 1200, -26{0}, 15{3}, 1, T7, 1, 1, &
54000{15}, 2, 30{3], , , 10
24500, Vehicles, -600, -14{0}, 20{3}, 1, T16, 1, 1


25800, Vehicles, 1200, -26{0}, 10{3}, 1, F26, 1, 1, &
54000{15}, 2, 30{3], , , 10
25800, Vehicles, -600, -14{0}, 20{3}, 1, T7, 1, 1


-1, 26500, Vehicles, 1200, -26{0}, 102{0}, 1, F3, 1, 1, &
54000{15}, 2, 30{3], , , 10

27800, Vehicles, 1200, -26{0}, 5{3}, 1, F4, 1, 1, &
32000{15}, 2, -10{3], , , 4, &
40000{15}, 2, 120{0], , , 10
27800, Vehicles, -600, -14{0}, 20{3}, 1, E9, 1, 1

30500, Vehicles, -600, -14{0}, 20{3}, 1, F26, 1, 1
30500, Vehicles, 1200, -26{0}, 5{3}, 1, S51, 1, 1
```

```
31700, Vehicles, -600, -14{0}, 20{3}, 1, F56, 1, 1
31700, Vehicles, 1200, -26{0}, -10{3}, 1, S10, 1, 1

32500, Vehicles, -600, -14{0}, 120{0}, 1, S5, 1, 1
32500, Vehicles, 1200, -26{0}, 112{0}, 1, S53, 1, 1

32800, Vehicles, -600, -14{0}, 120{3}, 1, S9, 1, 1
32800, Vehicles, 1200, -26{0}, 7{3}, 1, F30, 1, 1, &
54000{15}, 2, 120{0], , , 10

33100, Vehicles, 1500, 26{0}, 112{0}, 1, T15, 1, 1, &
54000{15}, 2, 120{0], , , 10
44978, Vehicles, 1500, 26{0}, 112{0}, 1, F56, 1, 1, &
54000{15}, 2, 120{0], , , 10




40100, Vehicles, -600, -14{0}, 120{0}, 1, F4, 1, 1
40400, Vehicles, 1200, -26{0}, 95{0}, 1, S5, 1, 1, &
54000{15}, 2, 120{0], , , 10




-1 cars on opposite side of road, --------------------------------------------------------------
-1, 0, Previously Defined Events, opposite.PDE


-1 crash
barriers************************************************************************
***********************************


-1,-----------------------------------------------------------------------------------------------


-1 middle left facing

0, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
3000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
6000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
9000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
12000, Static Object, 0, -0.1{0}, 0, 0, *, *,
```

```
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
15000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
18000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
21000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
24000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
27000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
30000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
33000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
36000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
39000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
42000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
45000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
48000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
51000, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10




0, Static Object, 0, -0.1{0}, 0, 0, *, *,
C:\Bosch\Projects\vico_2\autobahn\objects\buildings\leitplanke5.3DS, 3500,
10
1000, Tree, 55, 0, *1~3, 80{0}, 100{0}, 2
1000, Tree, 55, 0, *1~3, -80{0}, -100{0}, 2
```

Opposite.PDE

```
60, Vehicles, 1500,  26{0}, 120{0}, 3, T10, 1, 1
113, Vehicles, 1500, 14{0}, 120{0}, 3, S43, 1, 1
158, Vehicles, 1500, 14{0}, 120{0}, 3, T30, 1, 1
322, Vehicles, 1500, 26{0}, 120{0}, 3, S46, 1, 1
650, Vehicles, 1500, 36{0}, 120{0}, 3, T10, 1, 1
814, Vehicles, 1500, 26{0}, 120{0}, 3, T29, 1, 1
978, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1

1600, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
1900, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

2100, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
2500, Vehicles, 1500, 26{0}, 120{0}, 3, T10, 1, 1
2900, Vehicles, 1500, 36{0}, 120{0}, 3, T22, 1, 1

4000, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
4900, Vehicles, 1500, 36{0}, 120{0}, 3, T22, 1, 1


5200, Vehicles, 1500,  26{0}, 120{0}, 3, C16, 1, 1
5230, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
5450, Vehicles, 1500, 26{0}, 120{0}, 3, C1, 1, 1
5600, Vehicles, 1500, 36{0}, 120{0}, 3, C15, 1, 1
5714, Vehicles, 1500, 26{0}, 120{0}, 3, S9, 1, 1
5978, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1


6060, Vehicles, 1500,  26{0}, 120{0}, 3, T30, 1, 1
6113, Vehicles, 1500, 36{0}, 120{0}, 3, C1, 1, 1
6322, Vehicles, 1500, 26{0}, 120{0}, 3, S46, 1, 1
6650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
6978, Vehicles, 1500, 26{0}, 120{0}, 3, T30, 1, 1

7060, Vehicles, 1500,  26{0}, 120{0}, 3, T10, 1, 1
7113, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
7158, Vehicles, 1500, 26{0}, 120{0}, 3, S46, 1, 1
7322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
7650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
7978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

8060, Vehicles, 1500,  26{0}, 120{0}, 3, T10, 1, 1
8113, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
8158, Vehicles, 1500, 26{0}, 120{0}, 3, S46, 1, 1
8322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
8650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
8978, Vehicles, 1500, 26{0}, 120{0}, 3, S9, 1, 1

9000, Vehicles, 1500,  26{0}, 120{0}, 3, T10, 1, 1
9213, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
```

9322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
9700, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
9978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

10010, Vehicles, 1500, 26{0}, 120{0}, 3, T10, 1, 1
10213, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
10502, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
10600, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
10931, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

11060, Vehicles, 1500, 26{0}, 120{0}, 3, T10, 1, 1
11113, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
11322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
11650, Vehicles, 1500, 36{0}, 120{0}, 3, S35, 1, 1
11978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

12060, Vehicles, 1500, 26{0}, 120{0}, 3, T10, 1, 1
12158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
12322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
12650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
12978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

13060, Vehicles, 1500, 26{0}, 120{0}, 3, T10, 1, 1
13158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
13322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
13650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
13978, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1

14060, Vehicles, 1500, 26{0}, 120{0}, 3, T10, 1, 1
14158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
14322, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1
14650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
14978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

15060, Vehicles, 1500, 26{0}, 120{0}, 3, T2, 1, 1
15158, Vehicles, 1500, 36{0}, 120{0}, 3, T7, 1, 1
15322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
15650, Vehicles, 1500, 36{0}, 120{0}, 3, T29, 1, 1
15978, Vehicles, 1500, 26{0}, 120{0}, 3, C15, 1, 1

16060, Vehicles, 1500, 26{0}, 120{0}, 3, C16, 1, 1
16158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
16322, Vehicles, 1500, 26{0}, 120{0}, 3, S52, 1, 1
16650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
16978, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1

```
17060, Vehicles, 1500,  26{0}, 120{0}, 3, S43, 1, 1
17158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
17322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
17650, Vehicles, 1500, 36{0}, 120{0}, 3, S41, 1, 1
17978, Vehicles, 1500, 26{0}, 120{0}, 3, F4, 1, 1


18060, Vehicles, 1500,  26{0}, 120{0}, 3, T10, 1, 1
18158, Vehicles, 1500, 36{0}, 120{0}, 3, T18, 1, 1
18322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
18650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
18978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


19060, Vehicles, 1500,  26{0}, 120{0}, 3, F39, 1, 1
19158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
19322, Vehicles, 1500, 26{0}, 120{0}, 3, T19, 1, 1
19650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
19978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


20060, Vehicles, 1500,  26{0}, 120{0}, 3, F8, 1, 1
20158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
20322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
20650, Vehicles, 1500, 36{0}, 120{0}, 3, S44, 1, 1
20978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


21200, Vehicles, 1500,  26{0}, 120{0}, 3, T22, 1, 1
21400, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
21450, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
21600, Vehicles, 1500, 36{0}, 120{0}, 3, S9, 1, 1
21978, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1



22060, Vehicles, 1500,  26{0}, 120{0}, 3, S43, 1, 1
22158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
22322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
22650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
22978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


23060, Vehicles, 1500,  26{0}, 120{0}, 3, C1, 1, 1
23158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
23322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
23650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
23978, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1


24060, Vehicles, 1500,  26{0}, 120{0}, 3, T15, 1, 1
```

```
24158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
24322, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1
24650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
24978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


25060, Vehicles, 1500,  26{0}, 120{0}, 3, T22, 1, 1
25158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
25322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
25650, Vehicles, 1500, 36{0}, 120{0}, 3, T29, 1, 1
25978, Vehicles, 1500, 26{0}, 120{0}, 3, C15, 1, 1


25060, Vehicles, 1500,  26{0}, 120{0}, 3, T29, 1, 1
26158, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
26322, Vehicles, 1500, 26{0}, 120{0}, 3, S46, 1, 1
26650, Vehicles, 1500, 36{0}, 120{0}, 3, T10, 1, 1
26978, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1

27700, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
27900, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

27100, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
27600, Vehicles, 1500, 26{0}, 120{0}, 3, F4, 1, 1
27900, Vehicles, 1500, 36{0}, 120{0}, 3, T22, 1, 1

28000, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
28900, Vehicles, 1500, 36{0}, 120{0}, 3, T22, 1, 1


29200, Vehicles, 1500,  26{0}, 120{0}, 3, C16, 1, 1
29400, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
29450, Vehicles, 1500, 26{0}, 120{0}, 3, C1, 1, 1
29600, Vehicles, 1500, 36{0}, 120{0}, 3, C15, 1, 1
29978, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1


30060, Vehicles, 1500,  26{0}, 120{0}, 3, T30, 1, 1
30113, Vehicles, 1500, 36{0}, 120{0}, 3, C1, 1, 1
30322, Vehicles, 1500, 26{0}, 120{0}, 3, S46, 1, 1
30650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
30978, Vehicles, 1500, 26{0}, 120{0}, 3, T30, 1, 1

31060, Vehicles, 1500,  26{0}, 120{0}, 3, C1, 1, 1
31322, Vehicles, 1500, 26{0}, 120{0}, 3, F4, 1, 1
31650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
31978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

32060, Vehicles, 1500,  26{0}, 120{0}, 3, S9, 1, 1
32158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
32322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
32650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
```

```
32978, Vehicles, 1500, 26{0}, 120{0}, 3, S9, 1, 1

33000, Vehicles, 1500,  26{0}, 120{0}, 3, T30, 1, 1
33213, Vehicles, 1500, 36{0}, 120{0}, 3, T10, 1, 1
33322, Vehicles, 1500, 26{0}, 120{0}, 3, S9, 1, 1
33978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

34010, Vehicles, 1500,  26{0}, 120{0}, 3, F39, 1, 1
34458, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
34502, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
34931, Vehicles, 1500, 26{0}, 120{0}, 3, F45, 1, 1


35000, Vehicles, 1500,  26{0}, 120{0}, 3, S43, 1, 1
35322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
35700, Vehicles, 1500, 36{0}, 120{0}, 3, T10, 1, 1
35978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

36010, Vehicles, 1500,  26{0}, 120{0}, 3, T30, 1, 1
36512, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
36600, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
36931, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

37000, Vehicles, 1500,  26{0}, 120{0}, 3, E1, 1, 1
37322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
37700, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
37978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


38060, Vehicles, 1500,  26{0}, 120{0}, 3, T10, 1, 1
38158, Vehicles, 1500, 36{0}, 120{0}, 3, T18, 1, 1
38322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
38650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
38978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


39060, Vehicles, 1500,  26{0}, 120{0}, 3, F39, 1, 1
39158, Vehicles, 1500, 36{0}, 120{0}, 3, E2, 1, 1
39650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
39978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


40060, Vehicles, 1500,  26{0}, 120{0}, 3, F8, 1, 1
40158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
40322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
40650, Vehicles, 1500, 36{0}, 120{0}, 3, S44, 1, 1
40978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


41200, Vehicles, 1500,  26{0}, 120{0}, 3, F21, 1, 1
41400, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
41450, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
```

```
41600, Vehicles, 1500, 36{0}, 120{0}, 3, F19, 1, 1
41978, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1


42060, Vehicles, 1500,  26{0}, 120{0}, 3, F31, 1, 1
42158, Vehicles, 1500, 36{0}, 120{0}, 3, F42, 1, 1
42650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
42978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

43060, Vehicles, 1500,  26{0}, 120{0}, 3, C1, 1, 1
43158, Vehicles, 1500, 36{0}, 120{0}, 3, S34, 1, 1
43322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
43650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
43978, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1


44060, Vehicles, 1500,  26{0}, 120{0}, 3, F22, 1, 1
44158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
44322, Vehicles, 1500, 26{0}, 120{0}, 3, T15, 1, 1
44650, Vehicles, 1500, 36{0}, 120{0}, 3, E9, 1, 1
44978, Vehicles, 1500, 26{0}, 120{0}, 3, F56, 1, 1


45060, Vehicles, 1500,  26{0}, 120{0}, 3, T22, 1, 1
45158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
45322, Vehicles, 1500, 26{0}, 120{0}, 3, F7, 1, 1
45650, Vehicles, 1500, 36{0}, 120{0}, 3, T29, 1, 1
45978, Vehicles, 1500, 26{0}, 120{0}, 3, C15, 1, 1


46060, Vehicles, 1500,  26{0}, 120{0}, 3, T29, 1, 1
46158, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
46322, Vehicles, 1500, 26{0}, 120{0}, 3, S46, 1, 1
46650, Vehicles, 1500, 36{0}, 120{0}, 3, F9, 1, 1
46978, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1

47700, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
47900, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1

47100, Vehicles, 1500, 36{0}, 120{0}, 3, T30, 1, 1
47600, Vehicles, 1500, 26{0}, 120{0}, 3, F55, 1, 1
47900, Vehicles, 1500, 36{0}, 120{0}, 3, F8, 1, 1


48060, Vehicles, 1500,  26{0}, 120{0}, 3, T10, 1, 1
48158, Vehicles, 1500, 36{0}, 120{0}, 3, T18, 1, 1
48322, Vehicles, 1500, 26{0}, 120{0}, 3, F39, 1, 1
48650, Vehicles, 1500, 36{0}, 120{0}, 3, S43, 1, 1
48978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


49060, Vehicles, 1500,  26{0}, 120{0}, 3, F39, 1, 1
```

```
49158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
49322, Vehicles, 1500, 26{0}, 120{0}, 3, F38, 1, 1
49978, Vehicles, 1500, 26{0}, 120{0}, 3, F8, 1, 1


50060, Vehicles, 1500,  26{0}, 120{0}, 3, F8, 1, 1
50158, Vehicles, 1500, 36{0}, 120{0}, 3, S46, 1, 1
50322, Vehicles, 1500, 26{0}, 120{0}, 3, F41, 1, 1
50650, Vehicles, 1500, 36{0}, 120{0}, 3, S44, 1, 1
50978, Vehicles, 1500, 26{0}, 120{0}, 3, F3, 1, 1
```

Appendix B: Demographics Survey

Demographics Survey

* Required

1. **What is your full name?** *

2. **What is your email address?** *

3. **What's your date of birthday?**

   Example: December 15, 2012

4. **What's your gender?**

   Mark only one oval.

   - Male
   - Female
   - Other

5. **How many years have you been driving?**

6. **Do you drive?**

   Mark only one oval.

   - Once a year or less
   - Once a month
   - Once or more a week
   - Daily

7. **Do you own a car?**

   Mark only one oval.

   - Yes
   - No

8. **How many mileage do you drive a year?**

9. **Do you use ADAS or ACC? (select no if you are not sure)**
   Mark only one oval.
   - Yes
   - No

10. **Judge your driving skills (1 is best, 5 is amateur)**
    Mark only one oval.
    - 1
    - 2
    - 3
    - 4
    - 5

Experiment

11. **How easy was your first takeover?**
    Mark only one oval.
    - Difficult
    - Neutral
    - Easy

12. **How easy was your second takeover?**
    Mark only one oval.
    - Difficult
    - Neutral
    - Easy

13. **How easy was your third takeover?**
    Mark only one oval.
    - Difficult
    - Neutral
    - Easy

14. **How distracting was the email task? (5 is most distracting)**
Mark only one oval.

- 1
- 2
- 3
- 4
- 5

15. **How distracting was the Twenty Questions Task? (5 is most distracting)**
Mark only one oval.

- 1
- 2
- 3
- 4
- 5

16. **Please order these tasks according to their difficulty. The most difficult comes first**
Mark only one oval per row.

	Email	TQT
Rank #1		
Rank #2		

17. **Were you tired/exhausted/fatigued during the experiment? (5 is the highest, 1 is none)**
Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Appendix C: Consent Form and Information Sheet

Consent Form

How physiological behaviour of drivers correlates with their responses during takeover scenarios in highly automated driving

INFORMED CONSENT FORM

Taking Part **Please initial box**

The purpose and details of this study have been explained to me. I understand that this study is designed to further scientific knowledge and that all procedures have been approved by the Loughborough University Ethics Approvals (Human Participants) Sub-Committee.

I have read and understood the information sheet and this consent form.

I have had an opportunity to ask questions about my participation.

I understand that I am under no obligation to take part in the study, have the right to withdraw from this study at any stage for any reason, and will not be required to explain my reasons for withdrawing.

I agree to take part in this study. Taking part in the project will include being recorded (video); heart rate and eye movement data are collected.

Use of Information

I understand that all the personal information I provide will be treated in strict confidence and will be kept anonymous and confidential to the researchers unless (under the statutory obligations of the agencies which the researchers are working with), it is judged that confidentiality will have to be breached for the safety of the participant or others or for audit by regulatory authorities.

I understand that anonymised quotes may be used in publications, reports, web pages, and other research outputs. Video data will be strictly confidential and will be deleted by the end of the project.

I agree for the data I provide will be completely anonymous and shared publicly for other researchers. Video data will not be included in the shared data.

I agree to assign the copyright I hold in any materials related to this project to Mohamed Taher Alrefaie, the investigator of the experiment. ☐

Collected Data

I agree to get my hear rate, eye movement and to be video recorded during the experiment. ☐

__________________________ _______________________ _________
Name of participant [printed] Signature Date

_________________________ ______________________ _________
Researcher [printed] Signature Date

Information Sheet

How physiological behaviour of drivers correlate with their responses during takeover scenarios in highly automated driving

INFORMATION SHEET

Investigators Details:

Mohamed Taher Alrefaie
m.t.a.alrefaie@lboro.ac.uk
01509 223393
The School of Business and Economics, Loughborough LE11 3TU

Steve Summerskill
S.J.Summerskill2@lboro.ac.uk
01509222222
Loughborough Design School, Loughborough, LE11 3GR

Tom Jackson,
t.w.jackson@lboro.ac.uk
01509 223393
The School of Business and Economics, Loughborough LE11 3TU

Invitation

We would like to invite you to take part in our study. Before you decide we would like you to understand why the research is being done and what it would involve for you. One of our team will go through the information sheet with you and answer any questions you have. Talk to others about the study before making a decision if you wish.

What is the purpose of the study?

Fully autonomous vehicles will be a reality in 10-15 years. Until then, most vehicles will eventually require human intervention due to a system limitation or failure. But humans will be busy doing other activities such as reading, eating or checking their phone. These tasks cause a high mental workload and when vehicles requests a takeover, the driver may not be ready and hence will make a poor decision. In this study, we want to explore whether we can predict how good the human decision will be based on their eye movement, heart rate and their body posture right before a takeover is requested.

Who is doing this research and why?

Mohamed Taher Alrefaie will be the experimenter and main investigator of the experiment. He's a PhD student at Loughborough University and supervised by Tom Jackson and Steve Summerskill. The study is Part of Mohamed's PhD thesis.

Are there any exclusion criteria?

Participants who are below 18 or above 30 are excluded from this study. Also, participants with less than 2 years of driving experience are excluded.

What will I be asked to do?

You will be asked to come to Vehicles lab at Loughborough Design School. The experiment will last for a maximum of 90 minutes. You will be introduced to the experiment and equipment used. You will use a driving simulator to simulate driving and get used to it. The practice session will last for 20 minutes. Then, when ready the main experiment will start.

You will be asked to put on Tobii eyeglasses and Polar H7 heart rate monitor. Tobii Glasses use infrared to track eye movement and pupil size. There is a camera in front of Tobii Glasses to track what you see during the experiment. Polar H7 is a chest belt that tracks heart rate. These are two commercial products and are safe to use.

When the experiment starts, you will be asked to drive the car manually for a short period of time. The car will switch to automated mode. You may be asked to perform secondary tasks that simulate mental workload. These tasks include watching a video, reading news article, playing 20 questions task to guess an animal or writing an email. The chosen video will be in a general topic and won't require any prior knowledge. The news article will be in general business topics and should be easy to understand with no prior knowledge to the topic.

There will be four takeover requests that you need to handle. The automated system will encounter a problem in the road it cannot handle and will start a beep to alert you. After hearing the beep, you must look towards the road, understand the problem and take full driving control to handle the situation.

Once I take part, can I change my mind?

Yes. After you have read this information and asked any questions you may have if you are happy to participate we will ask you to complete an Informed Consent Form, however if at any time, before, during or after the sessions you wish to withdraw from the study please just contact the main investigator. You can withdraw at any time, for any reason and you will not be asked to explain your reasons for withdrawing.

However, once the results of the study are aggregated/published/dissertation has been submitted (expected to be by December 2018), it will not be possible to withdraw your individual data from the research.

Will I be required to attend any sessions and where will these be?

You will be required to be physically present in Vehicles Lab 0.16 at Loughborough Design School for up to 90 minutes.

How long will it take?

The experiment will take up to 90 minutes.

What personal information will be required from me?

We will ask for your name, gender, age, years of driving experience, car you usually drive, whether you used Adaptive Cruise Control (ACC) or not.

Are there any disadvantages or risks in participating?

The only major risk identified in this experiment is the high mental workload that may cause stress for few seconds. The experiment is done inside a driving simulator that is safe to use.

Will my taking part in this study be kept confidential?

You will be assigned an ID at the start of the experiment. Your name and ID will be kept in a separate file throughout the experiment and your name will not be included in any data processing.

On a separate file, your ID and your demographic information in addition to collected data (through devices) will be stored. This file will be processed and may be shared publicly on the internet.

Your video recording will NOT be shared with anyone beside the investigator (Mohamed Taher Alrefaie) of the study, his supervisors (Steve Summerskill, Tom Jackson), internal or external reviewers of the PhD study if needed. Once the study is over (expected December, 2018), video data will be deleted permanently.

Your data will be stored on the investigator's laptop or devices where he needs to conduct his study and research on throughout the duration of the study. These devices will be secured by a password and kept away from others.

I have some more questions; who should I contact?
You can ask the main investigator directly. If he doesn't answer or is unclear, please contact his supervisors.

What will happen to the results of the study?

Your results will be securely stored on a laptop, anonymised and processed. Some of these anonymous data will be shared publicly with other researchers.

The results will be published in the investigator's PhD thesis, conference papers or journal papers as well.

Anonymous data along with results will be publicly shared in a public repository for other researchers to work on.

What if I am not happy with how the research was conducted?

If you are not happy with how the research was conducted, please contact Ms Jackie Green, the Secretary for the University's Ethics Approvals (Human Participants) Sub-Committee:

Ms J Green, Research Office, Hazlerigg Building, Loughborough University, Epinal Way, Loughborough, LE11 3TU. Tel: 01509 222423. Email: J.A.Green@lboro.ac.uk

The University also has a policy relating to Research Misconduct and Whistle Blowing which is available online at http://www.lboro.ac.uk/committees/ethics-approvals-human-participants/additionalinformation/codesofpractice/ .

Is there anything I need to do before the sessions?

It's a driving experiment so please be prepared for a driving session. That means you should be fully awake, not tired, not under the influence of alcohol or drugs.

What are the possible benefits of participating?

You will have a good opportunity to experience driving future cars, see how your heart rate changes whilst driving and help researchers make future cars safer and easier to use.