

## Driver Psychology during Automated Platooning

Heikoop, Daniël

**DOI**

[10.4233/uuid:fe293c97-3f1e-4a39-bcbd-f5ab79d32d87](https://doi.org/10.4233/uuid:fe293c97-3f1e-4a39-bcbd-f5ab79d32d87)

**Publication date**

2017

**Document Version**

Final published version

**Citation (APA)**

Heikoop, D. (2017). *Driver Psychology during Automated Platooning*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:fe293c97-3f1e-4a39-bcbd-f5ab79d32d87>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

---

# **Driver Psychology during Automated Platooning**

Daniël David Heikoop

Driver Psychology during Automated Platooning

Proefschrift

ter verkrijging van de graad van doctor  
aan de Technische Universiteit Delft,  
op gezag van Rector Magnificus prof. ir. K. Ch. A. M. Luyben;  
Voorzitter van het College voor Promoties  
in het openbaar te verdedigen op  
woensdag 20 december 2017 om 15:00 uur

door

Daniël David HEIKOOP

Master of Science in Applied Cognitive Psychology, Universiteit Utrecht  
geboren in Leiderdorp, Nederland

This dissertation has been approved by the  
promoters: Prof. dr. ir. B. van Arem and Prof. dr. N. A. Stanton  
copromotor: Dr. ir. J. C. F. de Winter

Composition of the doctoral committee:

Rector Magnificus	chairperson
Prof. dr. ir. B. van Arem	Delft University of Technology
Prof. dr. N. A. Stanton	University of Southampton
Dr. ir. J. C. F. de Winter	Delft University of Technology

Independent members:

Prof. dr. M. Hagenzieker	Faculty of Civil Engineering and Geosciences, Delft University of Technology
Prof. dr. G. S. Matthews	Institute of Simulation and Training, University of Central Florida
Prof. dr. H. de Ridder	Faculty of Industrial Design Engineering, Delft University of Technology
Prof. dr. D. de Waard	Faculty of Behavioural and Social Sciences, Rijksuniversiteit Groningen

This research has been financially supported by the Marie Curie Initial Training Network (ITN) through the project Human Factors of Automated Driving (HFAuto), grant PITN-GA-2013-605817 to the project.

# Contents

Summary.....	viii
Samenvatting .....	xii
<b>Introduction .....</b>	<b>1</b>
The Rise of Automated Driving .....	2
Why Platooning? .....	3
Understanding Drivers' Psychological State during Platooning .....	4
Rationale for an Updated Psychological Model .....	5
The Aim of the Thesis .....	6
Methods for Human Factors of Automated Driving .....	7
Relevance of Human Factors Research in Automated Platooning.....	8
Thesis Contents.....	8
Chapter 2. Psychological Constructs in Driving Automation: A Consensus Model and Critical Comment on Construct Proliferation .....	9
Chapter 3. Effects of Platooning on Signal-Detection Performance, Workload, and Stress: A Driving Simulator Study .....	9
Chapter 4. Effects of Mental Demands on Situation Awareness during Automated Platooning: A Driving Simulator Study.....	10
Chapter 5. Acclimatizing to Automation: Driver Workload and Stress during Partially Automated Car Following in Real Traffic .....	11
Discussion.....	11
References .....	12
<b>Psychological Constructs in Driving Automation: A Consensus Model and Critical     Comment on Construct Proliferation .....</b>	<b>16</b>
Introduction .....	17
Method.....	21
Results .....	24
<i>Intermediate Steps between Psychological Constructs.....</i>	<i>27</i>
<i>Construct Proliferation.....</i>	<i>30</i>
Discussion.....	32
<i>Limitations of the Consensus Model.....</i>	<i>32</i>
<i>Characteristics of the Consensus Model.....</i>	<i>35</i>
<i>Recommendations for Future Research.....</i>	<i>36</i>
References .....	37
<b>Effects of Platooning on Signal-Detection Performance, Workload, Stress, and     Fatigue: A Driving Simulator Study.....</b>	<b>44</b>
Introduction .....	45
<i>Present Research .....</i>	<i>47</i>
Method.....	49
<i>Participants.....</i>	<i>49</i>
<i>Apparatus.....</i>	<i>50</i>
<i>Environment.....</i>	<i>50</i>
<i>Procedure.....</i>	<i>51</i>

<i>Conditions</i> .....	51
<i>Dependent Measures</i> .....	53
<i>Statistical Analyses</i> .....	56
Results .....	56
<i>Button Presses, Hits, and False Alarms</i> .....	56
<i>Self-Report Questionnaires: DSSQ and TLX</i> .....	57
<i>Heart Rate</i> .....	58
<i>Heart Rate Variability</i> .....	58
<i>Percentage Eyes Closed (PERCLOS)</i> .....	59
<i>Correlation Analysis</i> .....	59
Discussion.....	61
<i>Signal-Detection Performance</i> .....	63
<i>Self-Reported Stress and Workload</i> .....	65
<i>Heart Rate and Heart Rate Variability</i> .....	66
<i>Limitations</i> .....	66
Conclusion .....	69
References .....	70
<b>Effects of Mental Demands on Situation Awareness during Automated Platooning:</b>	
<b>A Driving Simulator Study</b> .....	<b>78</b>
Introduction .....	79
<i>Emergence of Automated Platooning</i> .....	79
<i>The Task of Drivers in a Platoon</i> .....	79
<i>Previous Research on Situation Awareness and Automated Driving</i> .....	79
<i>Aim of this Research</i> .....	81
Methods .....	82
<i>Participants</i> .....	82
<i>Apparatus</i> .....	83
<i>Environment</i> .....	83
<i>Procedure</i> .....	84
<i>Conditions</i> .....	85
<i>Dependent Measures</i> .....	86
<i>Statistical Analyses</i> .....	89
Results .....	89
<i>Self-Report Questionnaires: DSSQ and TLX</i> .....	89
<i>Performance on the 2-Back Task</i> .....	90
<i>Heart Rate</i> .....	91
<i>Heart Rate Variability</i> .....	92
<i>Eye Movements</i> .....	93
<i>Mental Models based on the Quick Association Check (QuACk)</i> .....	95
<i>Verbal Protocol Analysis</i> .....	97
Discussion.....	102
<i>Assessing the Effects of Mental Demands</i> .....	102
<i>Situation Awareness</i> .....	103
<i>Mental Models</i> .....	105
<i>Time-on-Task Effects</i> .....	107
<i>Measurement Issues</i> .....	107

<i>Further Research</i> .....	108
Conclusion .....	109
References .....	109
<b>Acclimatizing to Automation: Driver Workload and Stress during Partially Automated Car Following in Real Traffic</b> .....	<b>115</b>
Introduction .....	116
<i>Workload and Stress in Automated Driving</i> .....	116
<i>Prior Research on Workload and Stress in Automated Driving</i> .....	116
<i>Aim of the Present Study</i> .....	119
Methods .....	120
<i>Participants</i> .....	120
<i>Apparatus</i> .....	121
<i>Environment</i> .....	122
<i>Procedure</i> .....	122
<i>Independent Variables</i> .....	124
<i>Dependent Measures</i> .....	125
Results .....	130
Discussion .....	133
Conclusions and Recommendations .....	136
References .....	138
Supplementary materials .....	142
<b>Discussion and Conclusions</b> .....	<b>143</b>
Conclusions of the Thesis .....	144
Reflection on the Proposed Psychological Model .....	146
The Issue of Construct Proliferation .....	147
Simulator versus On-Road: A Comparison of Results .....	150
Validity of Psychophysiological Measures .....	154
Future Design and Implementations .....	155
References .....	161
<b>Acknowledgements</b> .....	<b>166</b>
<b>Curriculum Vitae</b> .....	<b>167</b>
<b>The 10 Propositions about my PhD Thesis</b> .....	<b>170</b>
<b>Appendices</b> .....	<b>171</b>
Appendix A: Experimental materials of Chapter 2 .....	171
<i>Appendix A.1: Reference list for Table 3</i> .....	171
Appendix B: Experimental Materials of Chapter 3 .....	174
<i>Appendix B.1: Experiment instruction sheet, Run 1, Condition No Task</i> .....	174
<i>Appendix B.2: Experiment instruction sheet, Run 2/3, Condition No Task</i> .....	176
<i>Appendix B.3: Experiment instruction sheet, Run 1, Condition Voluntary, Part: During             the experiment</i> .....	177
<i>Appendix B.4: Experiment instruction sheet, Run 2/3, Condition Voluntary</i> .....	178
<i>Appendix B.5: Experiment instruction sheet, Run 1, Condition Detection Task, Part:             During the experiment</i> .....	179

<i>Appendix B.6: Experiment instruction sheet, Run 2/3, Condition Detection Task</i>	180
<i>Appendix B.7: Participant information sheet</i>	181
<i>Appendix B.8: Consent form</i>	182
<i>Appendix B.9: ECG electrode placement information sheet</i>	183
<i>Appendix B.10: General information sheet</i>	184
<i>Appendix B.11: NASA-Task Load Index</i>	185
<i>Appendix B.12: Supplementary material: DSSQ 13 scales results, Conditionally</i>	186
<i>Appendix B.13: Supplementary material: DSSQ 13 scales results, Chronologically</i>	187
<b>Appendix C: Experimental Materials of Chapter 4</b>	188
<i>Appendix C.1: 2-back letter sequence score sheet</i>	188
<i>Appendix C.2: Experiment instructions sheet, Run 1, Condition Low Task Demands</i>	189
<i>Appendix C.3: Experiment instructions sheet, Run 2/3, Condition Low Task Demands</i>	190
<i>Appendix C.4: Experiment instructions sheet, Run 1, Condition Medium Task Demands, Part: During the experiment</i>	191
<i>Appendix C.5: Experiment instructions sheet, Run 2/3, Condition Medium Task Demands</i>	192
<i>Appendix C.6: Experiment instructions sheet, Run 1, Condition High Task Demands, Part: During the experiment</i>	193
<i>Appendix C.7: Experiment instructions sheet, Run 2/3, Condition High Task Demands</i>	194
<i>Appendix C.8: Participant information sheet</i>	195
<i>Appendix C.9: Consent form</i>	196
<i>Appendix C.10: ECG electrode placement information sheet</i>	197
<i>Appendix C.11: General information sheet</i>	198
<i>Appendix C.12: NASA-Task Load Index</i>	199
<b>Appendix D: Experimental Materials of Chapter 5</b>	200
<i>Appendix D.1: Participant information sheet</i>	200
<i>Appendix D.2: Participant information sheet</i>	201
<i>Appendix D.3: General information sheet</i>	202
<i>Appendix D.4: NASA-Task Load Index</i>	203
<i>Appendix D.5: Photographs of bridges 1-50 Northbound, retrieved from Google Maps</i>	204
<i>Appendix D.6: Photographs of bridges 1-47 Southbound, retrieved from Google Maps</i>	209



## Summary

With the rapid increase in vehicle automation technology, the call for understanding how humans behave while driving in an automated vehicle becomes more urgent. Vehicles that have automated systems such as Lane Keeping Assist (LKA) or Adaptive Cruise Control (ACC) not only support drivers in their journey, but also place them in a passive supervising role, scanning for potential hazardous stimuli in the environment or a system malfunction. More advanced technology that includes both lateral and longitudinal control and enables vehicles to drive at close distances from each other (called platooning technology) has the potential to reduce energy consumption and highway congestion. However, such technology places the driver in an even more critical position, as the time headway between vehicles is often below human reaction time (i.e., down to approximately 0.3 seconds). Little is known about driver behaviour, and the psychological constructs involved therewith, in automated platoons. This thesis investigates driver psychology during automated platooning.

The first objective of this thesis is to summarize what the literature states regarding the psychological constructs involved in automated driving. Based on a previously proposed model on psychological constructs in driving automation, a new model was developed that encompassed the psychological constructs used in the literature of automated driving as well as the interrelations between them (Chapter 2). During the search for literature, the issue of construct proliferation caused difficulty in formulating a model. Nevertheless, several well-established interrelationships between psychological constructs were identified, as well as some often overlooked mediating constructs.

The next step was to empirically test the developed model in an automated platooning environment. The first experiment (Chapter 3) assessed how signal detection performance, workload, and stress are influenced by automated platooning for extended periods of time in a driving simulator. Three conditions were compared between three 40 min drives in which the

## Summary

primary task was to pay attention to the road and intervene when a critical situation occurred. As a secondary task manipulation, a detection task condition in which the participant was tasked to detect red cars was compared with a voluntary condition, in which the participant was free to do as he/she wanted, and a no task condition. From the detection task's hit percentage it was clear that participants were able to remain relatively vigilant during the 40 min platooning drive. Results also showed that the voluntary condition was the least stressful, based on self-report questionnaires and psychophysiological measures. Moreover, a time-on-task effect was found, with the heart rate dropping significantly over the three runs that were performed. It was concluded that time-on-task effects are important to consider, and that the type-of-task during automated platooning has effects on the driver's psychological stress state.

The second experiment (Chapter 4) assessed the effects of mental demands on situation awareness during automated platooning. Next to that, it aimed to assess how mental models develop during extended exposure to non-critical automated platooning. Mental task demands were manipulated by means of a verbal 2-back task, creating a low, medium, and high task demand condition. Furthermore, by means of a "think-out-loud" protocol, participants' situation awareness was assessed. The results of the think-out-loud method showed that mental demands affected situation awareness, with an increase of mental demands resulting in a decrease in the occurrence of statements regarding situational features, such as "looking", and "overtaking". Moreover, time-on-task effects were found, as heart rate dropped and mental models grew over time. Based on these results and the psychological constructs assessed in this experiment, it was again concluded that time-on-task needs to have a prevalent role in Human Factors research regarding automated driving.

The third experiment (Chapter 5) involved an on-road experiment, and aimed to assess the generalizability of the results found in the simulator experiment as described in Chapter 3.

## Summary

With a Tesla Model S, participants were tasked to drive on the M40, M42, and M5 motorway, north- and southbound, with the Autopilot feature turned on, while following a lead vehicle at the closest time headway setting. In one of the two runs, participants were also asked to detect bridges, as a conceptual replication of the detection task performed in the simulator experiment. Similar to the results from the simulator experiment, a drop in heart rate occurred over time, and self-report scores on stress dimensions remained relatively constant between runs. The main conclusion from this experiment is that driving with the Autopilot on a highway involved a low level of self-reported workload and a heart rate which decreased over time.

The final chapter (Chapter 6) reflects on the proposed psychological model, and discusses the findings from the experiments, as well as the differences and similarities between on-road and simulator experimentation. First of all, the overall conclusion is that drivers are able to maintain a vigilant state for at least thirty minutes of automated platooning, although it must be noted that drivers are not able to remain perfectly vigilant. Moreover, the task to remain vigilant is mentally demanding and stressful to the driver, and makes them resort to non-driving tasks. One of the outcomes of this discussion is that the proposed model serves as a consensus model, rather than a truth model. Regarding the generalizability of results, it is concluded that several psychophysiological and self-reported results can be translated from a simulator experiment to an on-road environment. Furthermore, this chapter discusses the validity of psychophysiological measurements. It is argued that several of these issues arose due to the complex nature of human psychology, such that it is difficult to identify a 1-on-1 relationship between a physiological measure and a psychological construct. Also, the implications that the results have regarding the feasibility of automated platooning, and what design solutions could be offered to implement appropriate human-machine interfaces to ensure safe travel in an automated platoon are discussed. Automated platooning

## Summary

is deemed feasible if the appropriate applications, such as human-machine interfaces and driver-state monitors were implemented, so that the driver can drive the automated vehicle safely at all times. Otherwise, removal of the driver from driving responsibilities during automated platooning is advised.

## **Samenvatting**

Met de snelle toename in technologie dat automatisch rijden mogelijk maakt wordt de roep om begrip van menselijk gedrag in een automatisch rijdende auto steeds belangrijker. Auto's met automatische systemen zoals Lane Keeping Assist (LKA) of Adaptive Cruise Control (ACC) kunnen niet alleen de bestuurders ondersteunen in hun rit, maar kunnen hen ook in een passieve rol plaatsen waarin ze alleen uit dienen te kijken voor mogelijke gevaarlijke objecten en situaties of een systeem fout. Meer geavanceerde technologie dat zowel laterale als longitudinale controle combineert, en dat auto's de mogelijkheid biedt om dicht op elkaar te rijden (peloton technologie) heeft de potentie om energie verbruik en snelweg files te verminderen. Echter, zulke technologie plaatst de bestuurder in een nog kritischere positie, omdat de afstand tussen auto's vaak onder de menselijke reactiesnelheid ligt (dat is, tot aan ongeveer 0,3 seconden). Er is momenteel weinig bekend over het gedrag van een bestuurder, en de psychologische constructen die daarmee te maken hebben in een automatisch rijdend peloton. Deze these onderzoekt de psychologie van de bestuurder gedurende rit in een automatisch rijdend peloton.

De eerste taak van deze these is om samen te vatten wat de literatuur zegt over de psychologische constructen die te maken hebben met automatisch rijden. Gebaseerd op een eerder voorgesteld model over psychologische constructen in automatisch rijden is er een nieuw model ontwikkeld dat de huidige consensus in de literatuur bevat aangaande de psychologische constructen in automatisch rijden, en de relaties daartussen (Hoofdstuk 2). Tijdens de literatuurstudie leverde het probleem van construct proliferatie moeilijkheden op in het formuleren van een model. Desalniettemin waren er verscheidene welbekende relaties tussen psychologische constructen geïdentificeerd, alsmede enkele vaak over het hoofd geziene mediërende constructen.

De volgende stap was om het ontwikkelde model empirisch te testen in een omgeving van een automatisch rijdend peloton. Het eerste experiment (Hoofdstuk 3) trachtte te bepalen hoe de uitvoering van een detectietaak, en de hoeveelheid werkbelasting en stress beïnvloed worden door het automatisch rijden in een peloton voor lange tijd. Drie condities werden vergeleken met drie 40-minuten durende ritten waarin de primaire taak was om de aandacht op de weg te houden en in te grijpen wanneer een kritieke situatie zich voordeed. Als manipulatie van de secundaire taak werd een detectietaak, waarin de participant werd gevraagd om rode auto's te detecteren, vergeleken met een vrijwillige conditie, waarin de participant vrij was te doen wat hij/zij wilde, en met een conditie zonder taak. Vanuit het scoringspercentage van de detectie taak was het duidelijk dat participanten relatief waakzaam konden blijven gedurende de 40 minuten durende peloton rit. De resultaten lieten ook zien dat de vrijwillige conditie het minst stressvol was, gebaseerd op de zelf-rapportage vragenlijsten en psychofysiologische maten. Ook werd er een duidelijke “time-on-task” effect gevonden, omdat de hartslag significant verlaagde over de drie ritten die gedaan werden. Het was geconcludeerd dat “time-on-task” effecten belangrijk zijn om te overwegen, en dat de “type-of-task” gedurende automatisch rijden in een peloton van significant effect is op de bestuurder's psychologische stress staat.

Het tweede experiment (Hoofdstuk 4) bestudeerde de effecten van mentale vraag op situationeel bewustzijn gedurende automatisch rijden in een peloton. Daarnaast bestudeerde het hoe mentale modellen ontwikkelen gedurende verlengde blootstelling aan non-kritisch automatisch rijden in een peloton. In een lage, gemiddelde, en hoge taak vraag conditie, werd de mentale taak vraag gemanipuleerd door middel van een verbale 2-terug taak. Ook werd er door middel van een hardop-denken protocol het situationeel bewustzijn van participanten bestudeerd. De resultaten van het hardop-denken protocol lieten zien dat mentale vraag invloed had op situationeel bewustzijn, aangezien met toegevoegde vraag uitspraken over situationele

kenmerken, zoals “looking”, en “overtaking” minder frequent voorkwamen. Daarnaast werden ook “time-on-task” effecten gevonden, omdat de hartslag verlaagde, en mentale modellen groter werden over tijd. Gebaseerd op deze resultaten en de psychologische constructen die in dit experiment bestudeerd werden, was het wederom geconcludeerd dat “time-on-task” een belangrijke rol moet hebben in Human Factors onderzoek naar geautomatiseerd rijden.

Het derde experiment (Hoofdstuk 5) betrof een op-de-weg experiment en had als doel de generaliseerbaarheid van de gevonden resultaten in het simulator experiment te bepalen, zoals beschreven in Hoofdstuk 3. Participanten hadden de taak met een Tesla Model S te rijden op de M40, M42, en M5 snelweg, in noordelijke en zuidelijke richting, met de Autopilot aan, terwijl ze een volgauto volgden met de laagst mogelijke stand qua volgafstand. In één van de twee ritten werd aan de participanten ook gevraagd om bruggen te detecteren, als een conceptuele kopie van de detectietaak in het simulator experiment. Vergelijkbaar met de resultaten van het simulator experiment ontstond er een verlaging van de hartslag over tijd en relatief constante zelf-gerapporteerde scores op werkbelasting en stress dimensies. De hoofdconclusie uit dit experiment is dat rijden met de Autopilot op een snelweg een laag niveau aan zelf-gerapporteerde werkbelasting, en een dalende hartslag naar mate de tijd verstrijkt, inhoudt.

Het laatste hoofdstuk (Hoofdstuk 6) reflecteert op het voorgestelde psychologische model en bediscussieert de bevindingen van de experiment, alsmede de verschillen en overeenkomsten tussen op-de-weg en simulator experimenten. Allereerst is de algehele conclusie dat bestuurders een waakzame staat kunnen behouden voor minimaal dertig minuten in een automatisch rijdend peloton, met hierbij de kritische noot dat bestuurders niet perfect waakzaam kunnen blijven. Tevens is de taak om waakzaam te blijven mentaal belastend en stressvol voor de bestuurder, en laat hen neigen irrelevante taken uit te voeren.

Een van de uitkomsten van deze discussie is dat het voorgestelde model als een consensus model fungeert, in plaats van een waarheidsmodel. Aangaande de generaliseerbaarheid van de resultaten is het geconcludeerd dat verscheidene psychofysiologische en zelf-gerapporteerde resultaten vertaald kunnen worden van een simulator experiment naar een op-de-weg experiment. Daarnaast bediscussieert dit hoofdstuk de validiteit van psychofysiologische metingen. Het is gesuggereerd dat verscheidene problemen ontstonden door de complexe natuur van menselijke psychologie, zodanig dat het moeilijk is om een 1-op-1 relatie tussen een fysiologische maat en een psychologisch construct te identificeren. Daarnaast bediscussieert het de implicaties die de resultaten met zich meebrengen betreffende de huidige staat van automatisch rijdende pelotons, en welke ontwerp oplossingen aangeboden kunnen worden om toepasselijke mens-machine interfaces te kunnen implementeren zodanig dat veilig rijden in een automatisch rijdend peloton verzekerd kan worden. Automatisch rijden in een peloton wordt als haalbaar gezien als er passende toepassingen zoals mens-machine interfaces en systemen die de staat van een bestuurder monitoren worden geïmplementeerd, zodat de bestuurder altijd veilig in de automatische auto kan rijden. Anders wordt er geadviseerd om de bestuurder van de rijverantwoordelijkheden tijdens automatisch rijden in een peloton te ontdoen.



# 1

## Introduction

---

## **The Rise of Automated Driving**

Automated driving is a widely studied topic. Technological innovations such as Lane Keeping Assist (LKA) and Adaptive Cruise Control (ACC) are already available in consumer vehicles, allowing for automatic lateral and longitudinal control, respectively. Vehicles with this type of advanced driver assistance systems (ADAS) are considered to have SAE level 1 or 2 automation, meaning that the basic steering or acceleration/deceleration (level 1), or both simultaneously (level 2), can be done by the automated system, but also that the driver is still considered to “perform all remaining aspects of the dynamic driving task” (SAE International, 2016). One of the concepts that incorporates both lateral and longitudinal control of the vehicle is platooning, in which a group of vehicles drive closely together, with high coordination between each other, using technologies such as an Automated Highway System (AHS; Congress, 1994; Euler, 1990; Hancock et al., 1991) or Cooperative Adaptive Cruise Control (CACC; Van Arem, Van Driel, & Visser, 2006). More advanced technologies, in which the driver is permitted to engage in non-driving tasks (SAE levels 3–5), are currently demonstrated in various projects, such as AdaptIVe, Drive Me, PEGASUS, and SmartShuttle (Amditis, 2017; "Drive Me," 2017; "SmartShuttle," 2016; Zlocki, 2017).

With automated driving technologies being developed at a rapid pace, it becomes increasingly important to consider that the human driver plays a crucial role in their success (Kyriakidis et al., 2017). For example, if drivers dislike the automated driving system or misinterpret its capabilities, this may lead to a disuse of the technology and even cause unsafe situations (Beck, Dzindolet, & Pierce, 2007; Parasuraman & Riley, 1997; Saffarian, De Winter, & Happee, 2012). For instance, when an automated driving system sounds the alarm too late, the human driver will start to ignore the alarm (Abe & Richardson, 2006).

## Why Platooning?

A platoon consists of vehicles that drive closely together (e.g., a time headway of about 0.3 s; Ploeg, Van de Wouw, & Nijmeijer, 2014), typically using combined lateral and longitudinal control (Bergenheim et al., 2012; Ren & Green, 1994; see Figure 1.1 for a demonstration of an on-road platoon in The Netherlands and a simulated platoon). Compared to manual driving, platooning offers commercial and environmental benefits, such as improved fuel economy and traffic flow, and a decrease in carbon emission and insurance payments (Bergenheim, Huang, Benmimoun, & Robinson, 2010; Hochstädter & Cremer, 1997; Janssen, Zwijnenberg, Blankers, & De Kruijff, 2015; Karaaslan, Varaiya, & Walrand, 1991; Kunze et al., 2011; Tsugawa, Kato, & Aoki, 2011).

Truck platooning, with estimated benefits of up to more than €30k per truck per year (Janssen et al., 2015), has been the subject of various research projects. For example, the KONVOI, CHAUFFEUR I and II, and PATH projects have been demonstrating current technological advances (Fritz, Bonnet, Schiemenz, & Seeberger, 2004; Shladover et al., 1991; Wille, Röwenstrunk, & Debus, 2007), the Energy ITS and the European Truck Platooning Challenge projects have been assessing environmental benefits (Ellwanger & Wohlfarth, 2017; Tsugawa et al., 2011), and the European Truck Platooning Challenge and KONVOI projects have been conducting human factor studies (Ellwanger & Wohlfarth, 2017; Wille et al., 2007). But also platooning for regular cars has reached the interest of several research projects, such as PROMETHEUS, GCDC, and SARTRE, which aim to examine safety issues and to accelerate real-world implementation (Ploeg, Shladover, Nijmeijer, & Van de Wouw, 2011; "PROMETHEUS," 1986; Robinson, Chan, & Coelingh, 2010). Since the late 1980s, AHSs (Congress, 1994; Euler, 1990; Hancock et al., 1991) have gained interest among researchers, companies, and governmental bodies. Similarly, CACC uses vehicle-to-vehicle (V2V) communication to enable close car following without requiring segregated driving

lanes. Arguably, platooning and associated technologies such as CACC will become a reality in the foreseeable future (Janssen et al., 2015).



*Figure 1.1.* Demonstration of platooning in the Netherlands in 1998 ("Demo 98," 1998; Van Arem & Soeteman, 1998). Left = on-board view, Middle = drivers wave their hands out of the window to indicate that driving was hands-free. Right = Platooning in the driving simulator as used in the present thesis.

## Understanding Drivers' Psychological State during Platooning

Although automated platooning systems exist for several decades, the Human Factors issues associated with automated platooning are still relatively unexplored. With automated platooning systems becoming increasingly prevalent, the question arises how the drivers' psychological state will alter due to the role change from manual driving to (passively) supervising an automated driving system. A situation that might occur, albeit rarely, is hardware failure (see Seppelt & Victor, 2017). Also during the entering and exiting of a platoon, drivers may need to provide input and/or reclaim manual control (e.g., Levitan, Golembiewski, & Bloomfield, 1998; Nilsson, 2014). A reason for concern arises when drivers are subjected to long periods of automated driving, which is likely to occur in automated platooning. Classical psychological research has found that humans are not well able to remain attentive for prolonged periods of time, with a decrement in signal detection performance already occurring within 15 minutes (Mackworth, 1964).

In summary, it is vital to understand drivers' psychological state during prolonged monitoring of platooning, and to examine whether drivers are able to remain sufficiently alert.

The driver's state in a platoon will also affect the acceptance of the technology, and in turn, willingness to buy (Shladover, Campbell, Kailas, Boyd, & Torrey, 2015).

## **Rationale for an Updated Psychological Model**

Recently, research is concerned with how to get a driver back in the loop after a take-over request (TOR) provided by the automated vehicle. Amongst the topics that are under investigation regarding TORs are their effectiveness with respect to their modality (e.g., Naujoks, Mai, & Neukum, 2014; Petermeijer, Bazilinskyy, & De Winter, 2017; Walch et al., 2017), their temporal requirements (e.g., Eriksson & Stanton, 2017; Gold, Damböck, Lorenz, & Bengler, 2013; Lu, Coster, & De Winter, 2017), and what other aspects affect take-over quality (e.g., Bueno et al., 2016; Louw, Merat, & Jamson, 2015; Radlmayr, Gold, Lorenz, Farid, & Bengler, 2014). A selection of the recommendations drawn from such research is (1) to design for the 95<sup>th</sup> percentile, rather than the mean, because of large individual differences in, for instance, take-over time (Eriksson, Banks, & Stanton, 2017), (2) being distracted during automated driving significantly affects take-over quality (Louw et al., 2015), and (3) proper regeneration of the driver's situation awareness can take up to 20 seconds, which could have serious implications for the design of automated driving systems (Lu et al., 2017).

Little is currently known about how long it takes for drivers in automated vehicles to become distracted or disengaged from the driving task, or to lose their situation awareness. Moreover, what psychological constructs, such as workload or fatigue, contribute to that loss of engagement and situation awareness is currently still an unresolved topic. In order for, for instance, vehicle manufacturers to be able to anticipate on a driver's behaviour in an automated vehicle, it is important to understand the driver psychology. Therefore, it is suggested to model how psychological constructs interact in an automated driving domain (Michon, 1993; Stanton & Young, 2000). Stanton and Young (2000) proposed a

psychological model of automated driving in the 1990s, when automated driving technology was relatively new. Since then, it is likely that the knowledge on Human Factors of automated driving has been updated. Henceforth, this thesis starts with developing an updated model on the psychological constructs in automated driving, based on consensus within the current literature.

## **The Aim of the Thesis**

This thesis aims to answer what psychological constructs play a role in automated driving. More specifically, it aims to answer what happens to the driver's psychological state when he or she is being transported in a platoon of automated vehicles for a prolonged time. To measure and understand driver state is an important prerequisite for improving safety, because the problem domain needs to be understood before appropriate countermeasures (e.g., improved HMIs, adaptive automation) can be developed.

For instance, in an automated platoon, it is likely that a driver does not need to take over manual control for long periods of time, meaning that during large intervals no physical action is required from the driver. Therefore, it is expected that the state of the driver monitoring the automated platoon will digress towards a direction in which the driver is out of the loop for extended periods of time, which may have unwanted or unexpected consequences. If a driver of an automated platoon has to resort to (passively) monitoring for an automation failure, an obstacle on the road, or any other reason for a manual take-over, this could raise serious safety concerns (see also e.g., Körber, Cingel, Zimmermann, & Bengler, 2015; Louw, Madigan, Carsten, & Merat, 2016; Saxby, Matthews, Warm, Hitchcock, & Neubauer, 2013). Therefore, in an automated platoon, it is important to know and understand the driver state prior to a take-over request, to be able to take precautionary measures to get the driver back in a state in which he or she can respond appropriately to the given situation.

## **Methods for Human Factors of Automated Driving**

In research on Human Factors of automated driving, it is commonplace to use driving simulators, as simulators provide several benefits, such as relative ease of use, replicability of scenarios, and safety (Reed & Green, 1999). Nevertheless, on-road experimentation is sometimes preferred, for instance when it is expected that participants behave unrealistically in the simulator because of the lack of physical crash risk in the simulator (cf. Hallvig et al., 2013).

To investigate driver behaviour, three main methods are commonly used, namely task performance, psychophysiological measurement, and questionnaires (see e.g., Cain, 2007). Primary performance measures (such as steering- and braking quality and reaction times) are important, as they directly represent the effectiveness of, for example, a tested HMI (Cain, 2007). Next to primary task measures, a secondary task performance measure is important for measuring (or inducing) a driver's workload, because it inversely reflects primary task workload (Wickens, 1981). Physiological measurement is the second method of importance to Human Factors research (Cacioppo & Tassinary, 1990). The main advantage of using this type of measurement is its appropriateness for continuous and objective measurement (Cain, 2007). However, inferring a particular psychological state can be difficult, as these measures tend to be a general indicator for, for instance, global levels of stress or arousal, and are vulnerable for measuring multiple psychological constructs simultaneously (e.g., Cacioppo & Tassinary, 1990; Cain, 2007). The third method, conducting questionnaires, has the advantage to be practical, cheap, and easy to conduct. Moreover, it enables a researcher to collect large amounts of data in a relatively short period of time, and to measure a participant's private thoughts and feelings that no performance measure is able to capture. The drawbacks of the use of questionnaires is that they are filled out by the participant, so that it is impossible to tell whether what participants fill out is actually true. Moreover, what participants fill out is often

not in coherence with, for instance, performance measures when the limits of the participants' capabilities are examined (Yeh & Wickens, 1988). Because of the advantages and disadvantages of abovementioned three methods, this thesis applies a combination of these three methods in the experimental chapters.

## **Relevance of Human Factors Research in Automated Platooning**

As abovementioned, in order for automated driving to become a success, one needs to also have a user-centred focus, rather than merely an engineering focus. Consequential to this necessity is the understanding of the behaviour of the driver in such a vehicle, and the psychological constructs involved with establishing their behaviour (Michon, 1993; Stanton & Young, 2000). Moreover, by understanding driver psychology one can design appropriate HMIs and driver-state monitors to ensure safety and comfort of the driver. Thus, the topic addressed in this thesis (i.e., driver psychology in automated platooning) is not only of importance to expand our research knowledge base, but also to the industry for developing proper design strategies, and ultimately to legislation bodies for determining laws and rules to be adhered to in automated vehicles based on the psychological capabilities of humans driving in automated platoons.

## **Thesis Contents**

This thesis consists of four research papers. Each paper is self-contained, and is comprised of its own Introduction, Method, Results, Discussion, and References. The first paper is a literature study on the psychological constructs that are of relevance in automated driving. The three subsequent papers describe empirical research investigating selected psychological constructs during automated platooning / car following on a highway.



## **Chapter 2. Psychological Constructs in Driving Automation: A Consensus Model and Critical Comment on Construct Proliferation**

This paper reviews the literature on Human Factors of automated driving in order to revise a psychological model originally proposed by Stanton and Young (2000). More specifically, this paper identifies the most widely used psychological constructs in the literature, and describes the inter-relations between these constructs. Furthermore, the paper provides recommendations towards empirical research in order to test the relationships between the reviewed psychological constructs, and discusses the phenomenon of construct proliferation (i.e., the notion that some constructs appear to be so highly correlated that they may be practically indistinguishable). For instance, the psychological constructs “mental workload”, “attention”, and “situation awareness” are commonly mentioned in the literature. However, the mediating role of attention between mental workload and situation awareness can sometimes be missed, and the issue of construct proliferation provides a wide array of terms that can be interpreted as “mental workload”, ranging from “mental processing” to “cognitive activity”. The subsequent papers address selected constructs from this model.

## **Chapter 3. Effects of Platooning on Signal-Detection Performance, Workload, and Stress: A Driving Simulator Study**

This paper assesses the psychological effects of a visual monitoring task. Specifically, this paper examines the effects on driver workload and stress when drivers within an automated platoon have to remain vigilant by detecting salient stimuli on the road during the entire drive (Figure 1.2, left). Detection performance, and psychophysiological (i.e., cardiovascular and ocular) and self-report measures of stress and workload are measured, and the results and implications are discussed. The results point into the direction of a time-on-task effect based on a decline in heart rate, and a type-of-task effect based on that a voluntary condition yielded lower self-report scores on distress. Moreover, the fact that participants

were able to remain relatively vigilant raised questions regarding the applicability of classic vigilance literature in the domain of automated driving.

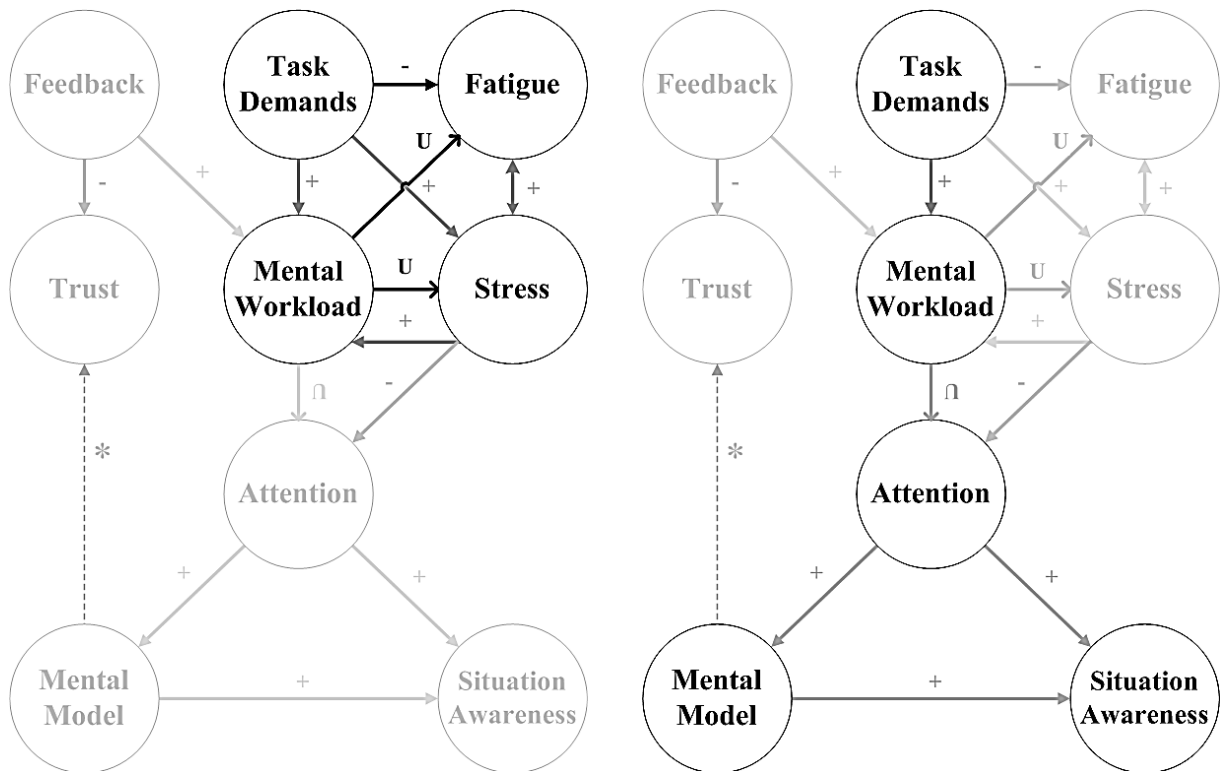


Figure 1.2. The psychological constructs and their inter-relationships investigated in Experiments 1 and 3 (Chapter 3 & 5; left), and Experiment 2 (Chapter 4; right).

## Chapter 4. Effects of Mental Demands on Situation Awareness during Automated Platooning: A Driving Simulator Study

Even when platooning drivers are visually attending the road (as studied in Chapter 3), they may still be mentally distracted. Accordingly, this paper assesses the effects of different levels of mental task demands on drivers' psychological state (Figure 1.2, right). In particular, by means of a working memory task, this paper induces three distinct levels of mental demand (low, medium, and high), and examines the effect on drivers' situation awareness by means of a verbal protocol method. Furthermore, this paper measures drivers' mental models at four moments during the experiment, and also uses, similar to the previous experiment, cardiovascular and eye tracking equipment to assess drivers' psychophysiological responses. Results suggest that added mental demands affect situation awareness, and that time-on-task

effects in automated platoon should be considered in Human Factors research on automated driving.

## **Chapter 5. Acclimatizing to Automation: Driver Workload and Stress during Partially Automated Car Following in Real Traffic**

The previous two experiments were conducted in a driving simulator. Driving simulators (see Figure 1.1, right) allow for controlled and safe measurements, but may pose questions about data validity. This paper seeks to investigate the generalizability of the effects found in the first experiment in a real-world setting. Accordingly, this paper measures workload and stress levels during an on-road car following experiment with an automated driving system. This on-road experiment uses the same psychophysiological and self-report measures as the previous two driving simulators studies, and includes an additional respiratory rate measure. Based on the results, it can be seen that an acclimatizing effect to the automated driving environment occurs.

## **Discussion**

The discussion of the thesis reflects on the psychological model described in the first paper, as well as on the results of three driving experiments. In particular, the key findings from the proposed psychological model, and its concurrent issue of construct proliferation is discussed. Also, the differences and similarities of simulator- and on-road experimentation, and the implications of the empirical results regarding the implementation of automated platooning in the real world are addressed. Finally, this thesis provides recommendations for future research.

## References

- Abe, G., & Richardson, J. (2006). The influence of alarm timing on driver response to collision warning systems following system failure. *Behaviour & Information Technology*, 25, 443-452. doi: 10.1080/01449290500167824
- AdaptIVe. (2014). from <https://www.adaptive-ip.eu/>
- Amditis, A. (2017). *European activities on connected and automated driving; The present and beyond – the adaptive and Autonet2030 use cases*. Paper presented at the Transportation Research Board 96th Annual Meeting, Washington, DC.
- Beck, H. P., Dzindolet, M. T., & Pierce, L. G. (2007). Automation usage decisions: Controlling intent and appraisal errors in a target detection task. *Human Factors*, 49, 429-437. doi: 10.1518/001872007X200076
- Bergenheim, C., Huang, Q., Benmimoun, A., & Robinson, T. (2010). Challenges of platooning on public motorways. *Proceedings of the 17th World Congress on Intelligent Transport Systems*, 1-12.
- Bergenheim, C., Pettersson, H., Coelingh, E., Englund, C., Shladover, S., & Tsugawa, S. (2012). Overview of platooning systems. *Proceedings of the 19th Intelligent Transportation Systems World Congress*.
- Bueno, M., Dogan, E., Hadj Selem, F., Monacelli, E., Boverie, S., & Guillaume, A. (2016). How different mental workload levels affect the take-over control after automated driving. *IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. doi: 10.1109/ITSC.2016.7795886
- Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45, 16-28. doi: 10.1037/0003-066X.45.1.16
- Cain, B. (2007). *A review of the mental workload literature*. (Report No. RTO-TR-HFM-121-Part-II), Defence Research and Development, Toronto, Canada.
- CHAUFFEUR I. (1996). from <http://vra-net.eu/wiki/index.php?title=Chauffeur>
- CHAUFFEUR II. (2000). from [http://vra-net.eu/wiki/index.php?title=Chauffeur II](http://vra-net.eu/wiki/index.php?title=Chauffeur_II)
- Congress, N. (1994). The automated highway system: An idea whose time has come. *Public Roads*, 58(1).
- Demo 98. (1998). from [https://www.dropbox.com/s/8hbgobncypbqmm7/Demo\\_98.rm?dl=0](https://www.dropbox.com/s/8hbgobncypbqmm7/Demo_98.rm?dl=0)
- Drive Me. (2017). from <http://www.volvocars.com/intl/about/our-innovation-brands/intellisafe/autonomous-driving/drive-me>
- Ellwanger, S., & Wohlfarth, E. (2017). Truck platooning application. *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*. Redondo Beach, CA.
- Energy ITS. (2008). from <http://vra-net.eu/wiki/index.php?title=Energy ITS - Automated Truck Platoon>
- Eriksson, A., Banks, V. A., & Stanton, N. A. (2017). Transition to manual: Comparing simulator with on-road control transitions. *Accident Analysis & Prevention*, 102, 227-234. doi: 10.1016/j.aap.2017.03.011
- Eriksson, A., & Stanton, N. A. (2017). Take-over time in highly automated vehicles: Non-critical transitions to and from manual control. *Human Factors*. doi: 10.1177/0018720816685832
- Euler, G. W. (1990). Intelligent Vehicle/Highway Systems: Definitions and applications. *ITE Journal*, 60(11), 17-22.
- Fritz, H., Bonnet, C., Schiemenz, H., & Seeberger, D. (2004). *Electronic tow-bar based platoon control of heavy duty trucks using vehicle-vehicle communication: Practical results of the CHAUFFEUR2 project*. Paper presented at the 11th World Congress on ITS, Nagoya, Japan.

- GCDC. (2016). from <http://www.gcdc.net/en/>
- Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57, 1938-1942. doi: 10.1177/1541931213571433
- Hallvig, D., Anund, A., Fors, C., Kecklund, G., Karlsson, J. G., Wahde, M., & Åkerstedt, T. (2013). Sleepy driving on the real road and in the simulator - A comparison. *Accident Analysis and Prevention*, 50, 44-50. doi: 10.1016/j.aap.2012.09.033
- Hancock, P. A., Caird, J. K., Johnson, S., Shehkar, S., Yang, A., Coyle, M., & Pawlacyk, L. (1991). Human factors safety issues in intelligent vehicle highway systems. *Safety News: Newsletter of the Safety Technical Group of the Human Factors Society*, 1-9.
- HFAuto. (2014). Human factors of automated driving., from <http://hf-auto.eu/>
- Hochstädter, A., & Cremer, M. (1997). Investigating the potential of convoy driving for congestion dispersion. *Proceedings of the IEEE Conference on Intelligent Transportation Systems*, 735-740. doi: 10.1109/ITSC.1997.660565
- Janssen, R., Zwijnenberg, H., Blankers, I., & De Kruijff, J. (2015). Truck platooning: Driving the future of transportation.
- Karaaslan, U., Varaiya, P., & Walrand, J. (1991). Two proposals to improve freeway traffic flow. *Proceedings of the American Control Conference* (pp. 2539-2544). Boston, MA.
- Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics*, 365-371. doi: 10.1016/j.promfg.2015.07.499
- Kunze, R., Haberstroh, M. A., Hauck, E., Ramakers, R., Henning, K., & Jeschke, S. (2011). Automated truck platoons on motorways – A contribution to the safety on roads. In S. Jeschke, I. Isenhardt & K. Henning (Eds.), *Automation, Communication and Cybernetics in Science and Engineering 2009/2010*: Springer: Berlin, Heidelberg.
- Kyriakidis, M., De Winter, J. C. F., Stanton, N. A., Bellet, T., Van Arem, B., Brookhuis, K., . . . Happee, R. (2017). A Human Factors perspective on automated driving. *Theoretical Issues in Ergonomics Science*. doi: 10.1080/1463922X.2017.1293187
- Leviton, L., Golembiewski, G., & Bloomfield, J. R. (1998). Human Factors issues for Automated Highway Systems. *ITS Journal – Intelligent Transportation Systems Journal: Technology, Planning, and Operations*, 4, 21-47. doi: 10.1080/10248079808903735
- Louw, T., Madigan, R., Carsten, O., & Merat, N. (2016). Were they in the loop during automated driving? Links between visual attention and crash potential. *Injury Prevention*. doi: 10.1136/injuryprev-2016-042155
- Louw, T., Merat, N., & Jamson, H. (2015). Engaging with highly automated driving: To be or not to be in the loop? *Proceedings of the 8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 190-196). Snowbird, Utah, USA.
- Lu, Z., Coster, X., & De Winter, J. C. F. (2017). How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving. *Applied Ergonomics*, 60, 293-304. doi: 10.1016/j.apergo.2016.12.003
- Mackworth, J. F. (1964). Performance decrement in vigilance, threshold, and high-speed perceptual motor tasks. *Canadian Journal of Psychology*, 18, 209-223. doi: 10.1037/h0083302
- Michon, J. A. (1993). *Generic intelligent driver support*. London: Taylor & Francis.
- Naujoks, F., Mai, C., & Neukum, A. (2014). The effect of urgency of take-over requests during highly automated driving under distraction conditions. In T. Ahram, W.

- Karwowski & T. Marek (Eds.), *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*. Kraków, Poland.
- Nilsson, J. (2014). *Safe transitions to manual driving from faulty automated driving system*. (Doctoral dissertation), Chalmers University of Technology, Gothenburg, Sweden.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 39, 230-253. doi: 10.1518/001872097778543886
- PATH. (1986). from <http://www.path.berkeley.edu/research/automated-and-connected-vehicles/truck-platooning>
- PEGASUS. (2017). from <http://www.pegasus-projekt.info/en/about-PEGASUS>
- Petermeijer, S. M., Bazilinsky, P., & De Winter, J. C. F. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied Ergonomics*, 62, 204-215. doi: 10.1016/j.apergo.2017.02.023
- Ploeg, J., Shladover, S. E., Nijmeijer, H., & Van de Wouw, N. (2011). Introduction to the special issue on the 2011 Grand Cooperative Driving Challenge. *IEEE Transactions on Intelligent Transportation Systems*, 13, 989-993. doi: 10.1109/TITS.2012.2210636
- Ploeg, J., Van de Wouw, N., & Nijmeijer, H. (2014). Lp string stability of cascaded systems: Application to vehicle platooning. *Proceedings of the IEEE Transactions on Control Systems Technology*, 22, 786-793. doi: 10.1109/TCST.2013.2258346
- PROMETHEUS. (1986). from <http://www.eurekanetwork.org/project/id/45>
- 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. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58, 2063-2067. doi: 10.1177/1541931214581434
- Reed, M. P., & Green, P. A. (1999). Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task. *Ergonomics*, 42, 1015-1037. doi: 10.1080/001401399185117
- Ren, W., & Green, D. (1994). Continuous platooning: A new evolutionary and operating concept for automated highway systems. *Proceedings of the American Control Conference*, 11(20), 21-25. doi: 10.1109/ACC.1994.751685
- Robinson, T., Chan, E., & Coelingh, E. (2010). Operating platoons on public motorways: An introduction to the SARTRE platooning programme. *Proceedings of the 17th World Congress on ITS*. Busan, South Korea.
- SAE International. (2016). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Washington, DC: SAE International.
- Saffarian, M., De Winter, J. C. F., & Happee, R. (2012). Automated driving: Human-factors issues and design solutions. *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting*, 2296-2300. doi: 10.1177/1071181312561483
- SARTRE. (2009). from <http://vra-net.eu/wiki/index.php?title=SARTRE>
- Saxby, D. J., Matthews, G. S., 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. *Journal of Experimental Psychology: Applied*, 19, 287-300. doi: 10.1037/a0034386
- Seppelt, B. D., & Victor, T. (2017). Potential solutions to Human Factors challenges in road vehicle automation. In G. Meyer & S. Beiker (Eds.), *Road Vehicle Automation 3: Lecture Notes in Mobility* (pp. 131-148).
- Shladover, S. E., Campbell, R., Kailas, A., Boyd, S., & Torrey, F. (2015). *Industry needs and opportunities for truck platooning*. UC Berkeley: California Partners for Advanced Transportation Technology.

- Shladover, S. E., Desoer, C. A., Hedrick, J. K., Tomizuka, M., Walrand, J., Zhang, W-B., . . . McKeown, N. (1991). Automatic vehicle control developments in the PATH program. *IEEE Transactions on Vehicular Technology*, 40, 114-130. doi: 10.1109/25.69979
- SmartShuttle. (2016). from <https://www.postauto.ch/en/smartshuttle>
- Stanton, N. A., & Young, M. S. (2000). A proposed psychological model of driving automation. *Theoretical Issues in Ergonomics Science*, 1, 315-331. doi: 10.1080/14639220052399131
- Tsugawa, S., Kato, S., & Aoki, K. (2011). An automated truck platoon for energy saving. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4109-4114. doi: 10.1109/IROS.2011.6094549
- Van Arem, B., & Soeteman, H. (1998). Highway to the future: A preview of the AVG demo '98. *Traffic Technology International*.
- Van Arem, B., Van Driel, C. J. G., & Visser, R. (2006). The impact of Cooperative Adaptive Cruise Control on traffic-flow characteristics. *IEEE Transactions on Intelligent Transportation Systems*, 7, 429-436. doi: 10.1109/TITS.2006.884615
- Walch, M., Mühl, K., Kraus, J., Stoll, T., Baumann, M., & Weber, M. (2017). From car-driver-handovers to cooperative interfaces: Visions for driver-vehicle interaction in automated driving. In G. Meixner & Müller C. (eds) (Eds.), *Automotive user interfaces. Human-computer interaction series* (pp. 273-294). Cham: Springer.
- Wickens, C. D. (1981). *Processing resources in attention and workload*. (Report No. EPL-81-3/ONR-81-3), Engineering-Psychology Research Laboratory University of Illinois, Champaign, IL.
- Wille, M., Röwenstrunk, M., & Debus, G. (2007). KONVOI: Electronically coupled truck convoys. In D. De Waard, F. O. Flemisch, B. Lorenz, H. Oberheid & K. A. Brookhuis (Eds.), *Human Factors for Assistance and Automation* (pp. 243-256). Maastricht, The Netherlands: Shaker Publishing.
- Yeh, Y-Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors*, 30(1), 111-120.
- Zlocki, A. (2017). *Securing automated driving – The database approach in PEGASUS*. Paper presented at the 1st European Conference on Connected and Automated Driving, Brussels, Belgium.

# Psychological Constructs in Driving Automation

---

## A Consensus Model and Critical Comment on Construct Proliferation

### Abstract

As automation in vehicles becomes more prevalent, the call for understanding the behaviour of the driver while driving an automated vehicle becomes more salient. Although a variety of driver behaviour models exist, and various psychological constructs have been said to be influenced by automation, an empirically testable psychological model of automated driving has yet to be developed. Building upon Stanton and Young's model of driving automation, this article presents an updated model of interrelated psychological constructs. The proposed model was created based upon a systematic literature search of driving automation papers and a subsequent quantification of the number of reported links between a selected set of psychological constructs. A secondary aim of this article is to reach consensus in the use of psychological constructs regarding driving automation. Henceforth special attention is paid to resolving the issue of construct proliferation.

Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2016). Psychological constructs in driving automation: A consensus model and critical comment on construct proliferation. *Theoretical Issues in Ergonomics Science*, 17(3), 284-303. doi:10.1080/1463922X.2015.1101507.



## **Introduction**

Automation in vehicles is becoming increasingly prevalent. Defined as the execution by a machine agent (usually a computer) of a function that was previously carried out by a human (Parasuraman & Riley, 1997), automation now has a major role in car driving. Present automation systems in cars range from ultrasonic or electromagnetic parking sensors that can inform the driver, to technology that can take over all longitudinal and lateral control tasks. Keeping in mind that the majority of vehicle accidents are caused by humans (e.g., Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006; Treat et al., 1977), it is argued that automation serves as a potential solution for driver safety.

Automating certain driver tasks might increase driver safety on the assumption that human operations are replaced with an infallible machine. However, due to risk compensation (also called behavioural adaptation; for overviews, see Elvik, 2006; Martens & Jenssen, 2012; Wilde, 1998), drivers may use automation in such a way that their behaviour changes (e.g., increasing speed, reducing headway, engaging in non-driving related tasks, etc.). Hence, more automation does not necessarily imply more safety. Another issue is that automation may result in mental overload in unforeseen circumstances (e.g., when automation fails and manual take-over is required), while mental underload is likely to occur during routine situations (Stanton & Marsden, 1996; Young & Stanton, 2002). A related concern is the lack of salient feedback from automation systems (Norman, 1990; Saffarian, De Winter, & Happee, 2012; Sarter, Woods, & Billings, 1997). Humans are notoriously bad at sustaining attention for prolonged periods of time while supervising rare signals (e.g., Mackworth, 1948; Molloy & Parasuraman, 1996), yet, paradoxically, this is exactly what drivers of automated vehicles are required to do. Such examples are just the tip of the iceberg when it comes to understanding human factors and ergonomics (HF/E) issues in automated driving. The field

would benefit from a model that can explain how drivers interact with their automated vehicles.

Thirty years ago, Michon (1985) performed a critical review of driver behaviour models, and distinguished four categories of models along two dimensions (behavioural vs. psychological, and taxonomic vs. functional). Michon argued that taxonomic-behavioural task analysis models do not account for the dynamic and complex environment of the driving task, and are therefore inadequate for modelling driver behaviour (Michon, 1985; Ranney, 1994). Alternatively, De Winter and Happee (2012) argued taxonomic-psychological (i.e., trait) models are a promising type of model, if developed through multivariate statistical approaches such as exploratory factor analysis (EFA) or structural equation modelling (SEM). Functional-behavioural models such as information flow control and adaptive control models focus on the physical motion of vehicles, without much consideration of drivers' motivations and cognitive processes. Hence, such models are less useful for understanding why drivers behave the way they do. Furthermore, most adaptive control models tend to be mathematically intricate, with limited generalizability (De Winter & Happee, 2012; see also Sheridan, 2004). The final category in Michon's (1985) overview contains functional-psychological models, focusing on driver motivation (e.g., Wilde, 1998) or cognitive processes (e.g., Bellet, Bailly-Asuni, Mayenobe, & Banet, 2009). Motivational models have a long history in the field of traffic psychology (Vaa, 2007), and describe the products of cognitive functions, such as beliefs and emotions (Michon, 1985). Although valuable, motivational models lack specificity and are therefore considered inadequate for modelling driver behaviour (De Winter & Happee, 2012). The cognitive process approach is considered by Michon (1985) to be an important approach in driver modelling, with the Adaptive Control of Thought-Rational (ACT-R) being one of the more popular methods (Anderson & Lebiere, 1998; Salvucci, 2006).

In order to understand and predict how people behave while driving automated vehicles, it is important to develop psychological models of driving with automation (Michon, 1993; Stanton & Young, 2000). Michon's (1985) categorization of driver behaviour models is concerned with non-automated vehicles. A few previous attempts have been made to develop psychological models for driving automation (e.g., Boer & Hoedemaeker, 1998; Stanton & Young, 2000), but despite rapid advances in vehicle automation technology and user uptake over the last two decades, an updated psychological model of driving automation is lacking.

Outside the domain of driving, several psychological models of human-automation interaction exist. Most of these models either describe automation psychology in general, thereby not addressing specific characteristics of the car driving task (e.g., Dzindolet, Pierce, Beck, Dawe, & Anderson, 2001; Parasuraman, Sheridan, & Wickens, 2000; Riley, 1989; Sanchez, 2009), and/or have a narrow scope as they address only a small number of psychological constructs (e.g., Endsley, 1995; Lee & See, 2004; Muir, 1994; Parasuraman & Manzey, 2010). For example, a model by Wickens, Lee, Liu, and Gordon-Becker (2004) qualitatively described the relationships between the psychological constructs trust and complacency in relation to automation use and automation reliability. This is a useful approach, but ignores other relevant constructs, such as mental workload and situation awareness.

This article is based upon a psychological model of driving automation developed by Stanton and Young (2000; Figure 2.1). When placing it into Michon's (1985) framework, the model of Stanton and Young (2000) can be categorised as a functional-psychological model, as it contains psychological constructs from both a cognitive (e.g., mental workload) and motivational (e.g., locus of control) perspective. One could also argue that it resembles a trait model approach, as their model also covers psychological constructs that were considered by some as psychological traits, for instance trust (e.g., Lee & See, 2004). The psychological

constructs of the Stanton and Young (2000) model were used as a basis for a literature search, with the aim of developing an updated psychological model of driving automation.

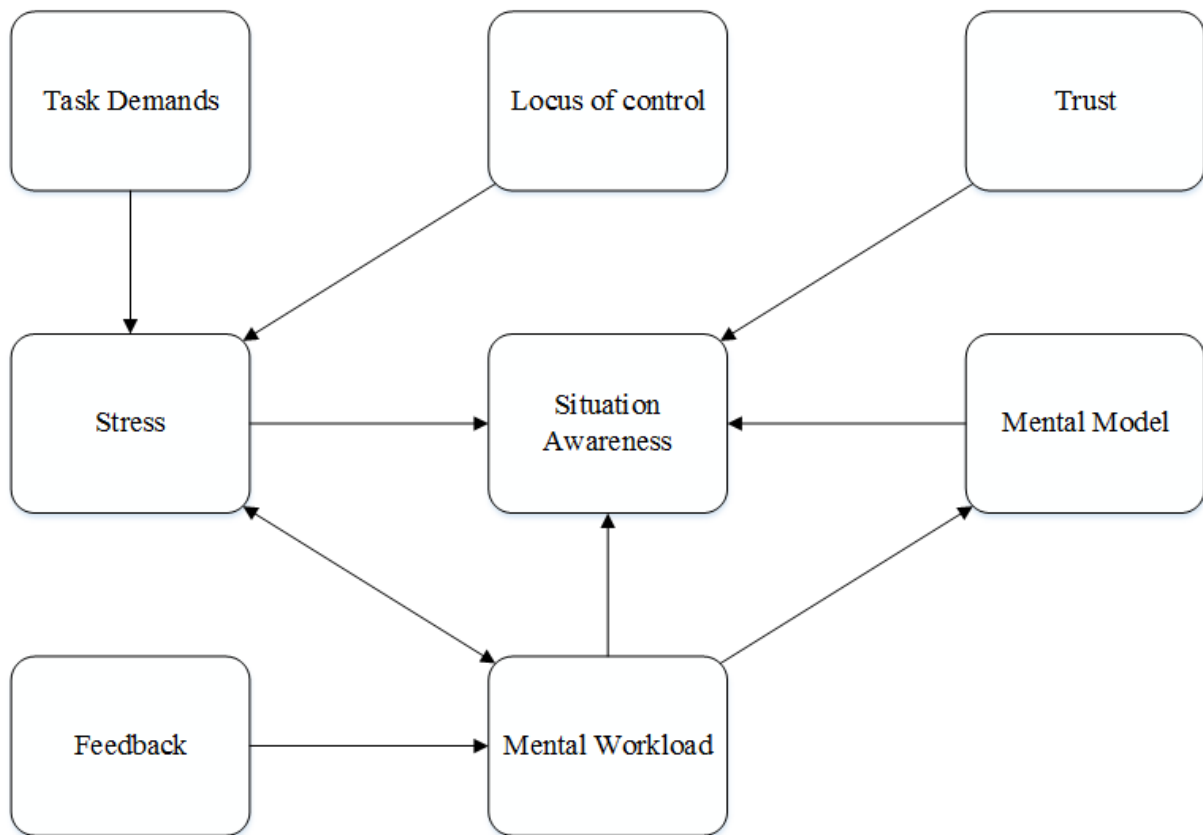


Figure 2.1. A proposed psychological model of driving automation (from Stanton & Young, 2000).

Stanton and Young (2000) used a theoretically oriented approach for the development of their model, using prominent literature of the time that presented results and discussions concerning the influences of automation on certain psychological constructs. Specifically, based on an earlier review by Stanton and Marsden (1996), Stanton and Young (2000) followed a deductive train of thought to identify psychological issues with vehicle automation, on which they built their psychological model of driving automation. Stanton and Young (2000) also provided a narrative review of these key psychological constructs, and correctly predicted that vehicular automation would have a major impact in the years to come. However, although Stanton and Young (2000) described the constructs used in their article in much detail with respect to their interrelationships, the model they proposed does not depict whether these interrelationships are causal or correlational, and whether the signs of the

effects are positive or negative. Implementation of this type of information would be a welcome addition to such a model.

The aim of the present review was to create an updated, testable version of the model developed by Stanton and Young (2000). The unique aspect of our approach is that it is descriptive and atheoretical. Meaning we measured how frequently key psychological constructs, as well as pairs of constructs, are reported in the scientific literature on automated driving. Based on this numeric information, we devised a model describing the interrelationships between the constructs. Hence, our approach offers a consensual description of the literature on the psychology of automated driving.

## **Method**

In their model (Figure 2.1), Stanton and Young (2000) used eight psychological constructs that were considered to have a critical impact upon behaviour when driving with automation: (1) Situation Awareness, (2) Mental Workload, (3) Mental Model, (4) Feedback, (5) Locus of Control, (6) Stress, (7) Task Demands and (8) Trust. The following seven extra constructs were also selected for the literature search: (9) Attention, (10) Vigilance, (11) Satisfaction, (12) Acceptance, (13) Arousal, (14) Complacency and (15) Fatigue. As opposed to the constructs used in the model of Stanton and Young (2000), which are primarily involved with short-term effects on driving psychology in automation, the seven extra psychological constructs serve to address the long-term effects (i.e., minutes to hours) on driving psychology in automation. We reasoned that the seven extra constructs are a welcome supplement, because it is likely that automated driving will be first deployed on highways (e.g., Bishop, 2005), where long-term use of automation is expected.

The eight constructs used in the Stanton and Young (2000), together with the seven supplementary constructs, were submitted to Google Scholar using Harzing's 'Publish or

Perish' software (version 4.10.1.5395; 8 October 2014). The use of Google Scholar over other academic search engines is advantageous, because Google Scholar is the only major academic search engine providing full text search (cf. Web of Science and Scopus, which only search abstracts). Furthermore, Google Scholar provides a substantially wider coverage of the scientific literature than other academic search engines (De Winter, Zadpoor, & Dodou, 2014; Gehanno, Rollin, & Darmoni, 2013; Shariff et al., 2013). As a result, Google Scholar also includes articles of lesser quality. To overcome the issue of literature quality a manual filtering method was applied and is described below.

Together with three domain-specific search terms (i.e., 'Driving Automation', 'Driver Automation' and 'Automated Driving') all possible unique combinations of the aforementioned constructs were used as search queries (a search query was for example: 'Driving Automation' AND 'Situation Awareness' AND 'Feedback'). With 3 domain-specific terms and 15 psychological constructs,  $3*((15*14)/2)$  unique combinations were possible, which yielded a total of 315 searches.

To make a distinction between articles of better quality and of lesser quality, a filter of a minimum of 10 citations per article was used as a threshold. These results were then filtered for duplicates, which resulted in a total of 224 unique articles containing any combination of two different constructs within the three domain-specific terms, henceforth referred to as driving automation. Patents and 'citations' (i.e., results that were displayed as either a patent or a citation in Google Scholar) were manually removed, as well as some obvious false positives (i.e., articles that were not about driving automation) and duplicates (i.e., articles that were dissimilar according to the results of the search tool, but after examination appeared to be similar).

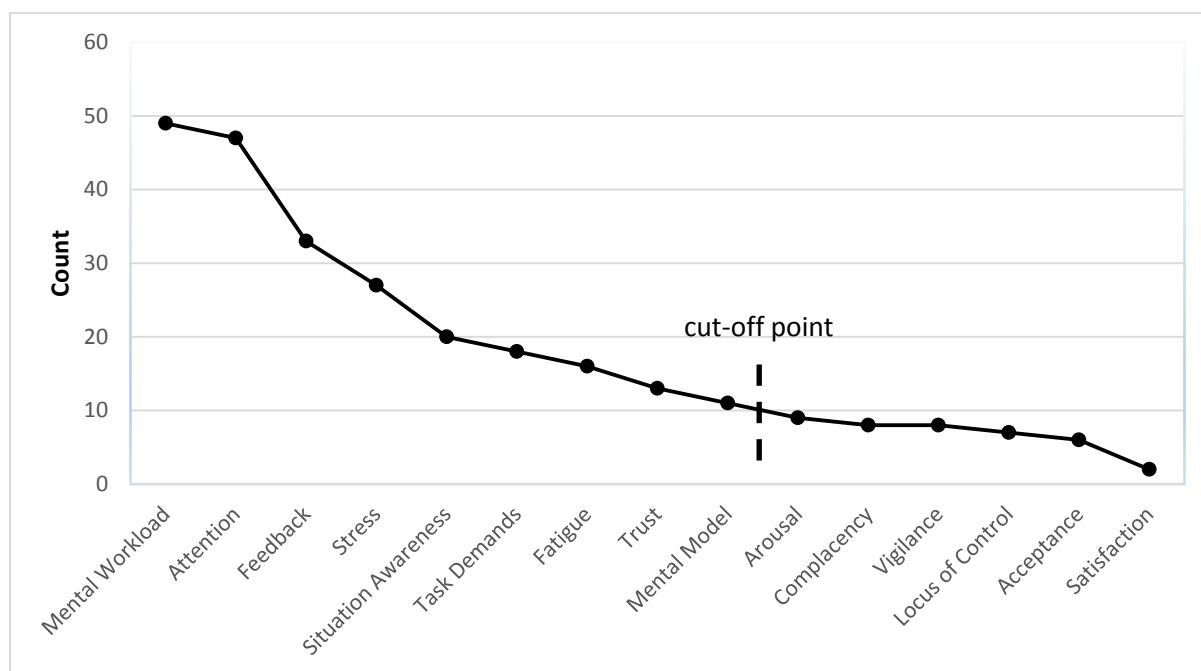
Of the resulting articles, the abstracts were read. Once an abstract of an article referred to either a link between constructs, or to an investigation of two or more constructs, thereby

showing a possibility of mentioning a link between constructs, the entire article was read. Whenever a link between constructs was mentioned, this link was noted down. A link could entail (1) the results of empirical or theoretical studies, (2) inferences made by authors based on previous (empirical or theoretical) studies, or (3) references to previous articles. This last option was seen as an acknowledgement of the existence and viability of this link by the author(s), thus reinforcing the link. The above process resulted in 43 unique articles mentioning a link between at least 2 of the 15 constructs.

In order to create an interpretable and parsimonious model, certain decisions had to be made as to whether or not to include each construct in the model. A simple counting scree plot (Figure 2.2) was used to assess the prevalence of each construct within the retrieved literature.

The constructs after the cut-off point (i.e., Arousal, Complacency, Vigilance, Locus of Control, Acceptance and Satisfaction) are henceforth left out of the model. A brief description of the psychological constructs used in our model is provided in Table 2.1. These definitions and descriptions were selected from the field of HF/E.

The construct with the widest variety of definitions was Mental Workload. The definition provided by Hart and Staveland (1988) is used in this article, as this definition applies best to the way the construct is being seen and used in this review, that is, Mental Workload being a human-centred construct rather than task-centred (Hart & Staveland, 1988). As for the construct of Feedback, one has to take into account that feedback is usually considered as automation-induced (e.g., visual or auditory signals), but can also be from the driving environment (e.g., seeing a car approaching, or hearing the engine is in the wrong gear). In this review, the articles used in the development of the model only refer to feedback as a form of automation-induced feedback.



*Figure 2.2.* Scree plot of the constructs used for the development of the psychological model of driving automation, derived from the 43 unique articles that resulted from the literature search. The count represents the number of times the construct was used in reference to a link with another construct viable for use in the model, that is, including a direction (which construct influences which construct?) and sign of the effect (does the one construct have a positive or negative relationship with the other construct?). The cut-off point was set at 10 counts.

## Results

The results of the literature search are shown in Table 2.2. The aim of this literature search was to determine causal links between constructs. Therefore, any mention of a link between constructs with a specific direction is thought to be causal. The constructs on the left hand side are the starters of a causal link towards the constructs on the top. This link is either a positive (1/0) or negative causal link (0/1). Table 2.2 does not show links mentioned or inferred in the articles without either a direction or a sign of effect. As this research is concerned with a psychological model of driving automation which can be empirically tested, it has a need for directions and signs of effect of links.

Table 2.2 clearly identifies two standout links; the positive causal links from Mental Workload and Task Demands to Attention (13 and 7 counts, respectively). Furthermore, Feedback appears to have a causal link towards Situation Awareness, Mental Workload, and



Attention. There does not seem to be a consensus on the sign of the effect, which can be explained by the type of Feedback (i.e., proper or improper) used in the research. For example, Coughlin, Reimer, and Mehler (2009) suggested that Feedback can be either distracting or attracting Attention depending on the form of Feedback.

Our psychological model on driving automation is shown in Figure 2.3, and consists of the constructs mentioned above, with their inferred relation to each other. During the development of this model certain inevitable subjective decisions had to be made, and various alternative hypotheses had to be considered. The detailed description of the steps that follows will help the reader understand and replicate the decisions made during this process.

Table 2.1

The nine psychological constructs used in the development of the updated model of driver psychology in automation, including definitions and references to key sources. In quotation marks are proper definitions given by their corresponding reference, whereas text without quotation marks are descriptions based on our inferred consensus in the literature.

Construct name	Definition/key description	Source
Situation Awareness	‘... the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future.’ ‘... knowing what’s going on so you can figure out what to do’	Endsley (1988) Adam (1993)
Mental Workload	A generally accepted definition does not exist, because of its applicability in a wide array of fields. ‘A hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance.’	Young, Brookhuis, Wickens, and Hancock (2014) Hart and Staveland (1988)
Mental Model	A mental model (MM) is a dynamic representation of the world. The concept dates back to 1943 where Craik mentioned a ‘small-scale model’ of the external reality. ‘In interaction with the environment, with others, and with the artefacts of technology, people form internal, mental models of themselves and of the things with which they are interacting.’	Johnson-Laird (1980) Norman (1983)

---

Feedback	A definition of feedback is usually dimensionalized, as it can be either quantitative or qualitative, as well as informational or motivational.	Norman (1990)
Stress	<p>‘Everyone knows what stress is, but nobody really knows.’</p> <p>‘Stress is the nonspecific response of the body to any demand made upon it.’</p> <p>In this field it is thought to be part of a group of conceptually similar constructs in which these are influenced by disturbances the traffic environment causes.</p>	<p>Selye (1973)</p> <p>Matthews (2002)</p>
Task Demands	‘The external demand, the goals that have to be reached, ...’	De Waard (1996)
Trust	Various approaches towards defining trust exist, ranging from it being an attitude or expectation, an intention or willingness to a behavioural result or vulnerable state.	Lee and See (2004)
Attention	‘There are two kinds of attention. Selective attention determines our ability to focus on certain sources of information and ignore others: ... Divided attention determines our ability to do more than one thing at once, ...’	Proctor and Van Zandt (2011)
Fatigue	‘Physiological fatigue is a loss of maximal force-generating capacity during muscular activity or a failure of the functional organ. ... Psychological fatigue ... has been defined as a state of weariness related to reduced motivation.’	Shen, Barbera, and Shapiro (2006)

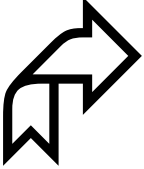
---

### **Intermediate Steps between Psychological Constructs**

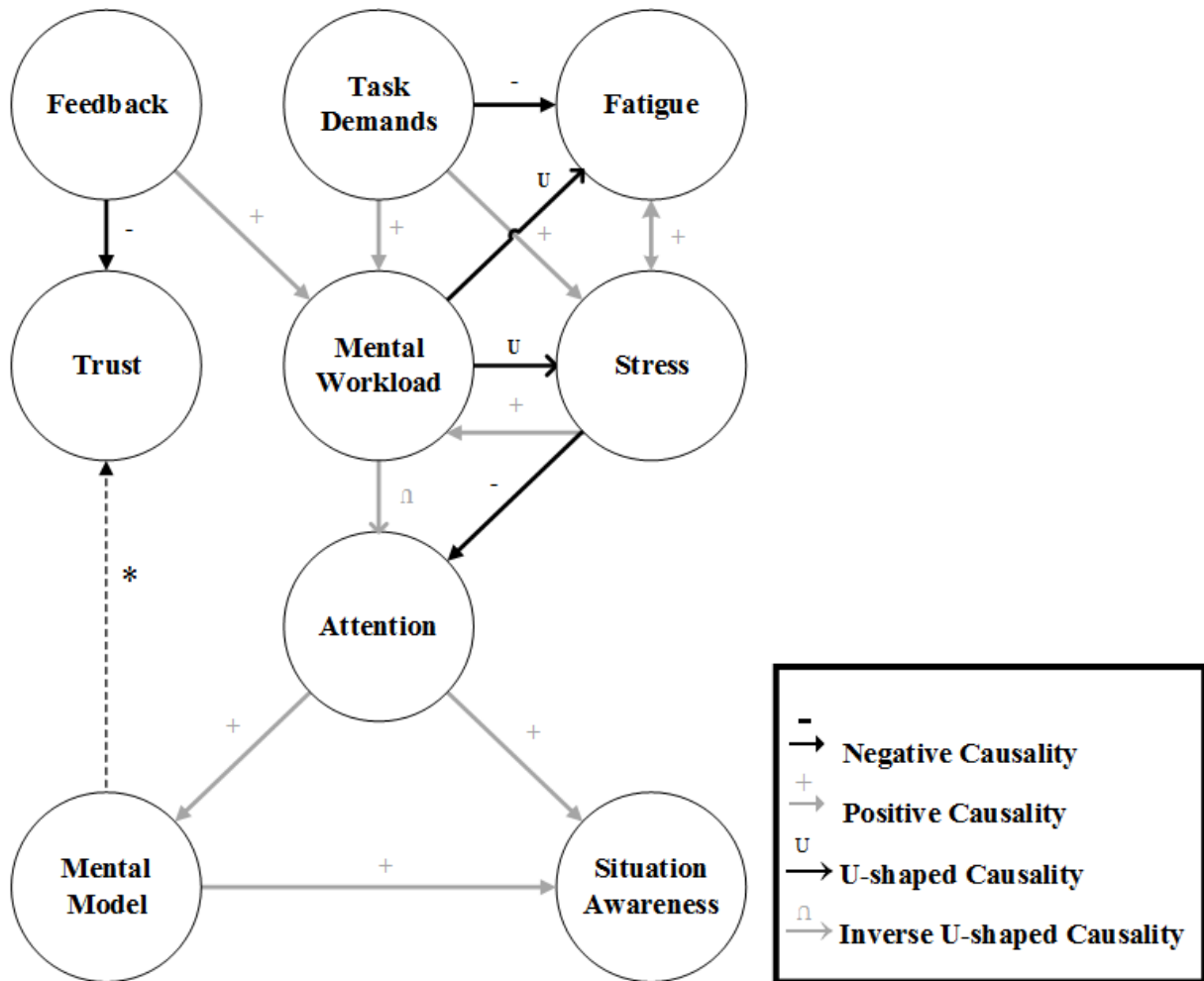
As many articles focus on different aspects of driver psychology, certain intermediate steps are prone to be overseen, resulting in a wide spread of interrelations. For example, one article may have mentioned a link between Situation Awareness and Mental Workload, and another article between Situation Awareness and Attention. Both articles may be correct, but one link may be an intermediate step within the other (e.g., Mental Workload → Attention → Situation Awareness). This aspect must be taken into consideration when developing a psychological model on driving automation. The results of this consideration are to be seen when comparing Table 2.2 with Figure 2.3. In the model shown in Figure 2.3, Task Demands are not directly linked to Attention, whereas Table 2.2 shows seven articles mentioning a direct positive causal link. However, after examination of these articles, most articles mentioned a task or construct that involves some form of Mental Workload in between Task Demands and Attention. As Mental Workload is composed of three elements, called (1) cognitive, (2) perceptual (i.e., visual, auditory, etc.), and (3) psychomotor workload (McCracken & Aldrich, 1984), tasks or constructs mentioned in those articles can often be considered as an element of Mental Workload. For example, Beede and Kass (2006) reported that using a phone and making decisions are examples of (cognitively) demanding tasks, whereas Collet, Guillot, and Petit (2010) referred to motor actions which can be seen as an aspect of (psychomotor) Mental Workload, and Stanley (2006) used a combination of both physical movement (looking over one's shoulder) and Mental Workload (remembering letters on a card). In summary, authors have often stated that they gave the participants a task (Task Demands) which in turn influenced Attention. What they did not explicitly mention is that these tasks were also cognitively/perceptually/psychomotorically demanding, as in requiring (some form of) Mental Workload. Hence, the conclusion was drawn that Task Demands influence Mental Workload, which in turn influences Attention.

Table 2.2

*Causation table between psychological constructs in driving automation. Each count represents an article mentioning a link between two constructs. The constructs on the left influence the constructs on the top either positively (1/0) or negatively (0/1).*

	Situation Awareness	Mental Workload	Mental Model	Feedback	Stress	Task Demands	Trust	Attention	Fatigue
Situation Awareness			1/0						
Mental Workload	2/4				5/0			13/1	2/3
Mental Model	3/0						2/0	1/0	
Feedback	2/2	3/4	1/0		0/2		0/2	3/5	0/1
Stress	0/1	1/0						0/3	
Task Demands		3/0			4/0			7/0	0/3
Trust	0/1				0/1			0/1	
Attention	4/0	3/0	1/0						
Fatigue					5/0			0/1	

Another feature of the model in Figure 2.3 is the absence of a direct link between Mental Workload and Situation Awareness. Stanton and Young (2005) concluded that Mental Workload influences Situation Awareness, but explained that Situation Awareness is comprised of various types of Attention. The relations between Attention, Situation Awareness, and Mental Workload have also been described in other articles (Matthews, 2002; Stanton & Young, 1998; Vahidi & Eskandarian, 2003). Furthermore, Stanton and Young (2000) discussed the construct Attention, but did not include it in their model.



*Figure 2.3.* The proposed updated psychological model of driving automation. It incorporates the interconnectedness of the constructs, similar to the model of Stanton and Young (2000), but adds a flow-like type to it for testability purposes. A black solid arrow indicates a negative causal link (e.g., when less Task Demands are being requested from the driver, more Fatigue is considered to be experienced by the driver). A grey solid arrow indicates a positive causal link. A black open arrow indicates a U-shaped causal link (e.g., both low and high Mental Workload evoke Stress), and, conversely, a grey open arrow indicates an inversed U-shaped causal link. \*The development of a Mental Model can recalibrate the trust towards the automation over time.

Similarly, a closer examination of the articles that mentioned a link between Feedback and Situation Awareness led us to believe no direct relation exists, but rather that Mental Workload acts as a mediator variable, which influences Attention to establish Situation Awareness, as explained by Stanton and Young (2005). This is also supported by (Endsley, 1995, 2000, 2015) who discussed how automation influences the attention requirements that are important for developing Situation Awareness.

Also, Stress and Fatigue have been represented differently in Figure 2.3 than in Table 2.2. This is largely due to the fact that Stress and Fatigue are very much related (Desmond & Matthews, 2009; Matthews, 2002; Saxby, Matthews, Hitchcock, & Warm, 2007) and therefore have similar characteristics (Desmond & Matthews, 2009).

### **Construct Proliferation**

Another issue encountered during the development of our driver model is the phenomenon of construct proliferation. That is, researchers may mistakenly assume that two highly correlated variables are unique constructs, while in reality they are one and the same construct. A review with examples on construct proliferation in psychology was performed by Schmidt (2010), in which he explained researchers often postulate new constructs, giving a false sense of differentiation. Ironically, the literature about this phenomenon suffers from construct proliferation itself, also coining terms such as construct redundancy (e.g., Le, Schmidt, Harter, & Lauver, 2010; Singh, 1991).

In the field of driving automation psychology the most prevalent occurrence of construct proliferation is with the construct Mental Workload. We encountered terms such as driver workload (Carsten & Nilsson, 2001; Flemisch et al., 2008; Funke, Matthews, Warm, & Emo, 2007; Schieben, Damböck, Kelsch, Rausch, & Flemisch, 2008), mental (over-/under-)load (Funke et al., 2007; Stanton, Dunoyer, & Leatherland, 2011; Young & Stanton, 1997), cognitive load (Carsten & Nilsson, 2001), or just simply over- (Coughlin et al., 2009; Matthews et al., 1999; Stanton et al., 2011; Young & Stanton, 1997) or underload (Navarro, Mars, & Young, 2011; Young & Stanton, 1997). This is just a selection of different uses which in fact can all be put under the same construct, in this article called Mental Workload (for further read, see McCracken & Aldrich, 1984). An overview of the varieties of psychological constructs used in this review is shown in Table 2.3.

Table 2.3

*List of psychological constructs from Figure 2.2 with corresponding varieties and references to their use. Corresponding references are shown in Appendix A.1.*

Original construct	Varieties	Authors
Situation Awareness	Situation Awareness	[3,6,18,24,25,33,34,35,36,37,38,40]
	Situational Awareness	[3]
	Awareness of the status	[20,32,38]
	Aware	[35]
Mental Workload	Mental Workload	[12,13,16,33,39,40,41,42]
	(Visual) Perception	[1]
	Mental/cognitive processing	[1]
	(Driver) work/mental/cognitive/over-/underload	[1,6,10,12,14,15,17,18,20,23,24,26,30,32,34,35,38,39,41,43]
	(Mental) demand(s)	[7,31,35]
	... decision making ...	[2,3,36]
	Cognitive resources	[3]
	Cognitive activity/task	[7,25]
	Listening to the radio (auditory)	[7]
	(demands on) Supervisory management	[18]
	Mentally capable	[21]
	Working memory	[23,35]
	Monitoring (demands)	[24,35,42]
Mental Model	Mental Model	[3,20,22,29,33,34,35,40]
	Mental representation	[3]
	Conceptual model	[17]
	Mental theory	[36]
	knowing how something behaves	[40]
Feedback	Feedback	[10,14,17,26,30,33,34,35,39,40]
	In-vehicle systems (e.g., ADAS)	[1,6,12,26,29,35,38,40]
	Provision of information	[5,22]
	Alerting methods (e.g., haptics)	[10]
	(False) Alarms/warnings	[17,20,26,27,31]
Locus of Control	of Locus of Control	[19,26,29,33,37,40]
	"Externals"/"Internals"	[34]
Stress	Stress	[5,10,12,15,23,24,26,32,33,34,35,40,41]
	Over arousal	[10]
	Distress	[12,15,23,24,43]
	Tenseness	[37]

Task Demands	Task demands	[15,23,24,33,35,39,41]
	Tasks that require ...	[1]
	... in demanding ... situations .../ ... demand(s)	[2,12,28,39]
	The ... task/ ... tasks	[9,11,18,23,31,32,34]
	Workload	[41]
Trust	Trust	[4,16,17,19,20,22,23,26,29,31,33,40]
Attention	Attention	[1,2,3,5,7,9,10,13,15,17,18,20,21,22,24,25,26,31,32,33,36,37,38,39,40,41,42]
	Attentional resources/processes	[3,9,16,17,35]
	(Distracting)	[10]
	(Attentional deficit)	[17]
	(Attentional shrinkage)	[26]
Vigilance	Vigilance	[5,7,8,18,25,28,33,35,39,41]
	Attentive	[34,37]
	(...)mobilization	[11]
Satisfaction	Satisfaction	[13,27,40]
	Acceptance	[27]
Acceptance	Acceptance	[5,10,13,14,17,20,27,31]
	Accepted	[13]
	Satisfaction	[27]
Arousal	Arousal	[7,8,10,18,28,33,41]
	Alerting stimulation	[10]
	Activation	[41]
Complacency	Complacency	[11,16,17,23,24,29,33]
	Abdicate	[4]
	Reliance	[6]
	Misplaced trust/overtrust	[40]
Fatigue	Fatigue	[11,12,15,18,23,24,26,28,32,34,35,42,43]

## Discussion

### Limitations of the Consensus Model

During the development of our model of driving automation, it has become clear that a descriptive literature-based approach is challenging for several reasons. First of all, during the literature search, the issue of construct proliferation arose. The development of the model required some subjective decisions, particularly due to the fact that the included articles had



different scopes. Hence, certain assessments had to be made concerning interpretation of terms and phrases shown in Table 2.3. The authors acknowledge that the phenomenon of construct proliferation could have led to misinterpretation of results or inferences made in this review.

A second limitation is that we set a cut-off mark at 10 counts of links (Figure 2.2), which may be seen as a somewhat arbitrary decision. However, we argue that in any statistical model a trade-off has to be made between model complexity (i.e., number of constructs included) and model comprehensiveness (i.e., representativeness of the literature). This situation is the same in EFA where deciding on the number of factors to retain is essentially a model selection problem (Preacher, Zhang, Kim, & Mels, 2013). There has been considerable debate in the statistical literature about how to interpret scree plots and how to create appropriate cut-off criteria (e.g., Fabrigar, Wegener, MacCallum, & Strahan, 1999). In summary, the cut-off mark we selected should be seen as an appropriate solution, not necessarily the only valid solution.

As a third limitation, we acknowledge that the interpretations made in this review of the links inferred in the articles are subject to the risk of inferring causality from correlation. However, the emphasis within the review was on developing a model based on causal relations. Therefore, only relationships that were regarded as causal have been used for the development of the model, whereas correlations have been disregarded, as correlation does not imply causality. The links used for the model were based on either empirical, theoretical, or deductive evidence. Some of the relationships were more ambiguously described than others, which increased the risk of misinterpretation.

A final limitation is that consensus does not imply ‘truth’ or ‘most appropriate’, because what researchers do in their experiments and report in their articles may be much influenced by the availability heuristic. For example, it is possible that some constructs are

highly used by researchers for the simple reason that others have also used them, thereby contributing to self-reinforcing behaviour, or the ‘Matthew effect’ (Merton, 1968). It is also possible that some constructs are used more than others for the reason that they can be more easily observed and measured (Dekker & Hollnagel, 2004). For example, Stress can be determined using a user-friendly self-report questionnaire such as the Dundee Stress State Questionnaire (DSSQ; Matthews et al., 1999), whereas other types of constructs, such as Attention, may require the application of cumbersome and expensive techniques, such as eye-tracking systems. Furthermore, since driving automation is a relatively new domain, newly arising constructs might be underrepresented compared to well-established constructs. Relatedly, the 10-citations threshold may be disputed as a means of filtering article quality. It is acknowledged that this threshold discriminates against recent and poorly accessible articles (e.g., articles that are not available online, or articles that are not published in open-access journals). However, recently published articles are not yet well-established, so by definition they do not belong in a consensual model.

Despite these limitations, the clear strength of our review is its descriptive ‘consensual’ approach, rather than it being based on personal theorizing. In other words, the exact aim of our research was to determine how psychological constructs are used, and not to determine how they ought to be used. Our approach to this review may be regarded as unusual in the sense of theory building and critical analysis. However, the insights that our atheoretical model brings are a complement to the many existing theoretically oriented driver models (e.g., Michon, 1985; Ranney, 1994), and might therefore be important to the scientific community. Furthermore, the validity of the updated model is supported by the fact that it contains certain well-known relationships, such as Mental Workload and Task Demands having a positive causal link with Attention. With this in mind, it should be repeated that the

updated model is not a definite statement, but rather a proposal based on what the literature infers regarding the intricate psychology of driving an automated vehicle.

### **Characteristics of the Consensus Model**

Despite the issues mentioned above it was possible to develop an updated psychological model of driving automation that provides interesting insights as well as new directions towards future research. For instance, although the Stanton and Young (2000) model was proposed fifteen years ago and not based on a systematic literature search, their model and ours show a high degree of similarity regarding their structure. Although many similarities appear, our updated model contains more detail concerning the interrelations of the psychological constructs within the model. Primarily, the implementation of directions and signs of effect to the links between the psychological constructs in the model might serve as a welcome addition in comparison to the Stanton and Young (2000) model, as this gives direction to future empirical research as to how these psychological constructs relate to each other.

Many types of psychological models of driving behaviour exist (Michon, 1985). The same applies to psychological models of automation (e.g., Dzindolet et al., 2001; Endsley, 1995; Lee & See, 2004; Muir, 1994; Parasuraman & Manzey, 2010; Parasuraman et al., 2000; Riley, 1989; Wickens et al., 2004). However, models that link psychological constructs in the driving automation domain are rare. Stanton and Young (2000) proposed a psychological model of driving automation based on the existing literature, a model that formed the foundations of the present article. By means of a thorough literature review, we developed a new model of driving automation (Figure 2.3), in an attempt to validate and expand the proposed model of Stanton and Young (2000). Our review attempted to describe a general consensus amongst researchers, concerning the interplay between psychological constructs in driving automation. However, it was not expected that the results, definitions and use of the

constructs would be as widespread as they appear to be. In light of the issue of construct proliferation, it must be stressed, as was previously by Schmidt (2010) that the use and misuse of constructs may lead to inconsistent or even false data. Future research could try to tackle the problem of construct proliferation, for example by assessing convergent and divergent validity using the multitrait-multimethod matrix approach (Campbell & Fiske, 1959).

### **Recommendations for Future Research**

The model proposed in this review was developed based on existing literature and hence not empirically tested. Furthermore, this model was designed to fit specifically within the domain of driving automation. An extension of the model with important non-psychological constructs (e.g., ability, authority) and legal issues such as responsibility might increase its applicability.

Observable measures such as performance, behaviour, and safety (i.e., incidents and accidents) were not included in the model either. We argue that the psychological status of the driver is at the root of the causal tree. By this we mean that a driver's psychological status in the automated vehicle (e.g., the driver's level of Situation Awareness) will likely determine how effectively he/she will perform in a take-over scenario, which in turn determines the risk of collisions. We decided to not model such effects, but believe that our model could in principle be extended and thereby become a potential predictor of performance, behaviour, and safety.

Additionally, future research might investigate whether the model can be adjusted for application to other domains, such as aviation. However, caution should be taken with this approach. For example, the field of aviation (with professional pilots performing highly procedural work) is fundamentally different from the automotive domain (with high degrees of freedom for drivers who usually have received only basic training; see e.g., Wheeler & Trigs, 1996).

In an attempt to test our newly proposed psychological model of driving automation, each part of the model could be assessed through empirical investigations. These investigations may take the form of using driving simulator technology with self-report questionnaires and psychophysiological measurements, such as eye trackers (e.g., Jamson, Merat, Carsten, & Lai, 2013; Merat, Jamson, Lai, & Carsten, 2012) or heart rate measurements (e.g., Brookhuis & De Waard, 2010), and assess two or more of the psychological constructs to determine their interrelations.

For example, it is possible to change the level of automation (in the likes of Jamson et al., 2013) or to alter the characteristics of a platoon (cf. Skottke, Debus, Wang, & Huestegge, 2014), and to investigate the impact this has on the psychological constructs. Banks, Stanton, and Harvey (2014) emphasised several concerns regarding the changing role of the driver due to increasing amounts of subsystems within automated vehicles. These authors pointed out that driving automation may entail a change or increase of the monitoring environment for the driver. Hence, it is acknowledged that not only the degree of automation, but also the number of subsystems to interact with, is an important subject to take into account.

Experiments could also investigate the impact of different types of secondary tasks (e.g., verbal vs. spatial tasks; see Young & Stanton, 2007) on the correlations between a set of psychological constructs. Proposals to variations in experimental manipulations are thus plentiful, but one cannot ignore variations in individual differences, which may be as plentiful and important to understand (Cronbach, 1957).

## References

- Adam, E. C. (1993). Fighter cockpits of the future. *Proceedings of the 12th DASC., AIAA/IEEE Digital Avionics Systems Conference*, 318-323. doi: 10.1109/DASC.1993.283529
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

- Bach, K. M., Jæger, M. G., Skov, M. B., & Thomassen, N. G. (2009). Interacting with in-vehicle systems: Understanding, measuring, and evaluating attention. *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, 453-462. doi: 10.1145/1671011.1671070
- Banks, V. A., Stanton, N. A., & Harvey, C. (2014). Sub-systems on the road to vehicle automation: Hands and feet free but not 'mind' free driving. *Safety Science*, 62, 505-514. doi: 10.1016/j.ssci.2013.10.014
- Beede, K. E., & Kass, S. J. (2006). Engrossed in conversation: The impact of cell phones on simulated driving performance. *Accident Analysis & Prevention*, 38, 415-421. doi: 10.1016/j.aap.2005.10.015
- Bellet, T., Bailly-Asuni, B., Mayenobe, P., & Banet, A. (2009). A theoretical and methodological framework for studying and modelling drivers' mental representations. *Safety Science*, 47, 1205-1221. doi: 10.1016/j.ssci.2009.03.014
- Bishop, R. (2005). Intelligent vehicle R&D: A review and contrast of programs worldwide and emerging trends. *Annales Des Télécommunications*, 60, 228-263. doi: 10.1007/BF03219820
- Boer, E. R., & Hoedemaeker, M. (1998). Modeling driver behavior with different degrees of automation: A hierarchical decision framework of interacting mental models. *Proceedings of the 17th European Annual Conference on Human Decision Making and Manual Control*, 63-72.
- Brookhuis, K. A., & De Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis and Prevention*, 42, 898-903. doi: 10.1016/j.aap.2009.06.001
- Brookhuis, K. A., De Waard, D., & Jansen, W. H. (2001). Behavioural impacts of advanced driver assistance systems-An overview. *European Journal of Transport and Infrastructure Research*, 1, 245-253.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation using the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105. doi: 10.1037/h0046016
- Carsten, O. M. J., & Nilsson, L. (2001). Safety assessment of driver assistance systems. *European Journal of Transport and Infrastructure Research*, 1, 225-243.
- Collet, C., Guillot, A., & Petit, C. (2010). Phoning while driving II: A review of driving conditions influence. *Ergonomics*, 53, 602-616. doi: 10.1080/00140131003769092
- Collet, C., Petit, C., Priez, A., & Dittmar, A. (2005). Stroop color-word test, arousal, electrodermal activity and performance in a critical driving situation. *Biological Psychology*, 69, 195-203. doi: 10.1016/j.biopsycho.2004.07.003
- Coughlin, J. F., Reimer, B., & Mehler, B. (2009). *Driver wellness, safety & the development of an awarecar*. Cambridge, MA: MIT AgeLab.
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684. doi: 10.1037/h0043943
- De Waard, D. (1996). *The measurement of drivers mental workload*. (Doctoral dissertation), Rijksuniversiteit Groningen.
- De Winter, J. C. F., & Happee, R. (2012). Modelling driver behaviour: A rationale for multivariate statistics. *Theoretical Issues in Ergonomics Science*, 13, 528-545. doi: 10.1080/1463922X.2010.546437
- De Winter, J. C. F., Zadpoor, A. A., & Dodou, D. (2014). The expansion of google scholar versus web of science: A longitudinal study. *Scientometrics*, 98, 1547-1565. doi: 10.1007/s11192-013-1089-2
- Dekker, S., & Hollnagel, E. (2004). Human factors and folk models. *Cognition, Technology & Work*, 6, 79-86. doi: 10.1007/s10111-003-0136-9

- Desmond, P. A., Hancock, P. A., & Monette, J. L. (1998). Fatigue and automation-induced impairments in simulated driving performance. *Transportation Research Record*, 1628, 8-14. doi: 10.3141/1628-02
- Desmond, P. A., & Matthews, G. S. (2009). Individual differences in stress and fatigue in two field studies of driving. *Transportation Research Part F*, 12, 265-276. doi: 10.1016/j.trf.2008.12.006
- Dzindolet, M. T., Pierce, L. G., Beck, H. P., Dawe, L. A., & Anderson, B. W. (2001). Predicting misuse and disuse of combat identification systems. *Military Psychology*, 13, 147-164. doi: 10.1207/S15327876MP1303\_2
- Elvik, R. (2006). To what extent can theory account for the findings of road safety evaluation studies? *Accident Analysis & Prevention*, 36, 841-849. doi: 10.1016/j.aap.2003.08.003
- Endsley, M. R. (1988). Situation Awareness Global Assessment Technique (SAGAT). *Proceedings of the National Aerospace and Electronics Conference* (pp. 789-795): New York: IEEE.
- Endsley, M. R. (1995). Towards a theory of situation awareness in dynamic systems. *Human Factors*, 37, 32-64. doi: 10.1518/001872095779049543
- Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 1-24). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Endsley, M. R. (2015). Final reflections: Situation awareness models and measures. *Journal of Cognitive Engineering and Decision Making*, 9, 101-111. doi: 10.1177/1555343415573911
- Engström, J., & Hollnagel, E. (2007). A general conceptual framework for modelling behavioural effects of driver support functions. In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments: Critical issues in driver interactions with intelligent transport systems* (pp. 61-84). London: Springer.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-299. doi: 10.1037//1082-989X.4.3.272
- Flemisch, F. O., Kelsch, J., Löper, C., Schieben, A., Schindler, J., & Heesen, M. (2008). *Cooperative control and active interfaces for vehicle assistance and automation*. Paper presented at the FISITA World Automotive Congress, Munich, Germany.
- Funke, G., Matthews, G. S., Warm, J. S., & Emo, A. K. (2007). Vehicle automation: A remedy for driver stress? *Ergonomics*, 50, 1302-1323. doi: 10.1080/00140130701318830
- Gehanno, J-F., Rollin, L., & Darmoni, S. (2013). Is the use of google scholar enough to be used alone for systematic reviews. *BMC Medical Informatics and Decision Making*, 13(7), 1-5. doi: 10.1186/1472-6947-13-7
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139-183. doi: 10.1016/S0166-4115(08)62386-9
- Hoc, J-M., Mars, F., Milleville-Pennel, I., Jolly, É., Netto, M., & Blosseville, J-M. (2006). Human-machine cooperation in car driving for lateral safety: Delegation and mutual control. *Le Travail Humain*, 69, 153-182. doi: 10.3917/th.692.0153
- Hoc, J-M., Young, M. S., & Blosseville, J-M. (2009). Cooperation between drivers and automation: Implications for safety. *Theoretical Issues in Ergonomics Science*, 10, 135-160. doi: 10.1080/14639220802368856
- 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. *Transportation Research Part C*, 30, 116-125. doi: 10.1016/j.trc.2013.02.008

- Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, 4, 71-115. doi: 10.1207/s15516709cog0401\_4
- Kazi, T. A., Stanton, N. A., Walker, G. H., & Young, M. S. (2007). Designer driving: Drivers' conceptual models and level of trust in adaptive cruise control. *International Journal of Vehicle Design*, 45, 339-360. doi: 10.1504/IJVD.2007.014909
- Kim, Y-S. (2001). *Effects of driver, vehicle, and environment characteristics on collision warning systems design*. (Master's thesis), Linköping Institute of Technology.
- Klauer, S. G., Dingus, D. R., Neale, T. A., Sudweeks, J., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic study data*. (Report No. DOT HS 810 594), National Highway Traffic Safety Administration, Washington, DC.
- Landau, K. (2002). Usability criteria for intelligent driver assistance systems. *Theoretical Issues in Ergonomics Science*, 3, 330-345. doi: 10.1080/14639220210137748
- Le, H., Schmidt, F. L., Harter, J. K., & Lauver, K. J. (2010). The problem of empirical redundancy of constructs in organizational research: An empirical investigation. *Organizational Behavior and Human Decision Processes*, 112, 112-125. doi: 10.1016/j.obhdp.2010.02.003
- Lee, J., & See, K. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46, 55-80. doi: 10.1518/hfes.46.1.50\_30392
- Ma, R., & Kaber, D. B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics*, 35, 939-953. doi: 10.1016/j.ergon.2005.04.002
- Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1, 6-21. doi: 10.1080/17470214808416738
- Martens, M. H., & Jenssen, G. D. (2012). Behavioural adaptation and acceptance. In A. Eskandarian (Ed.), *Handbook of Intelligent Vehicles* (pp. 117-138). London: Springer.
- Matthews, G. S. (2002). Towards a transactional ergonomics for driver stress and fatigue. *Theoretical Issues in Ergonomics Science*, 3, 195-211. doi: 10.1080/14639220210124120
- Matthews, G. S., Campbell, S. E., Desmond, P. A., Huggins, J., Falconer, S., & Joyner, L. A. (1999). Assessment of task-induced state change: Stress, fatigue and workload components. In M. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- McCracken, J. H., & Aldrich, T. B. (1984). *Analyses of selected LHX mission functions: Implications for operator workload and system automation goals*. (Report No. ASI 479-024-84(B)), Fort Rucker, AL: Anacapa Sciences Inc.
- Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. M. J. (2012). Highly automated driving, secondary task performance and driver state. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 54, 762-771. doi: 10.1177/0018720812442087
- Merton, R. K. (1968). The Matthew Effect in science. *Science*, 159(3810), 56-63. doi: 10.1126/science.159.3810.56
- Michon, J. A. (1985). A critical view of driver behavior models: What do we know, what should we do? In L. Evans & R. Schwing (Eds.), *Human Behavior and Traffic Safety* (pp. 485-520). New York, NY: Plenum Press.
- Michon, J. A. (1993). *Generic intelligent driver support*. London: Taylor & Francis.



- Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 38, 311-322. doi: 10.1177/001872089606380211
- Muir, B. M. (1994). Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics*, 37, 1905-1922. doi: 10.1080/00140139408964957
- Navarro, J., Mars, F., & Young, M. S. (2011). Lateral control assistance in car driving: Classification, review and future prospects. *IET Intelligent Transport Systems*, 5, 207-220. doi: 10.1049/iet-its.2010.0087
- Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental Models* (pp. 7-14). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Norman, D. A. (1990). The problem with automation: Inappropriate feedback and interaction, not over-automation. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 327(1241), 585-593. doi: 10.1098/rstb.1990.0101
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52, 381-410. doi: 10.1177/0018720810376055
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 39, 230-253. doi: 10.1518/001872097778543886
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, 30, 286-297. doi: 10.1109/3468.844354
- Piao, J., & McDonald, M. (2008). Advanced driver assistance systems from autonomous to cooperative approach. *Transport Reviews: A Transnational Transdisciplinary Journal*, 28, 659-684. doi: 10.1080/01441640801987825
- Preacher, K. J., Zhang, G., Kim, C., & Mels, G. (2013). Choosing the optimal number of factors in exploratory factor analysis: A model selection perspective. *Multivariate Behavioral Research*, 48, 28-56. doi: 10.1080/00273171.2012.710386
- Proctor, R. W., & Van Zandt, T. (2011). *Human Factors in Simple and Complex Systems* (2nd ed.). Boca Raton, FL: CRC Press.
- Ranney, T. A. (1994). Models of driving behavior: A review of their evolution. *Accident Analysis & Prevention*, 16, 733-750. doi: 10.1016/0001-4575(94)90051-5
- Rauch, N., Kaussner, A., Krüger, H-P., Boverie, S., & Flemisch, F. O. (2009). The importance of driver state assessment within highly automated vehicles. *Proceedings of the ITS – 16th World Congress and Exhibition on Intelligent Transport Systems and Services*.
- Riley, V. (1989). A general model of mixed-initiative human-machine systems. *Proceedings of the Human Factors and Ergonomics Society 33rd Annual Meeting*, 124-128. doi: 10.1177/154193128903300227
- Rudin-Brown, C. M. (2010). 'Intelligent' in-vehicle intelligent transport systems: Limiting behavioural adaptation through adaptive design. *IET Intelligent Transport Systems*, 4, 252-261. doi: 10.1049/iet-its.2009.0151
- Saffarian, M., De Winter, J. C. F., & Happee, R. (2012). Automated driving: Human-factors issues and design solutions. *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting*, 2296-2300. doi: 10.1177/1071181312561483
- Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 48, 362-280. doi: 10.1518/001872006777724417

- Sanchez, J. (2009). Conceptual model of human-automation interaction. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting*, 1403-1407. doi: 10.1177/154193120905301850
- Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. In G. Salvendy (Ed.), *Handbook of Human Factors & Ergonomics* (2nd ed., pp. 1926-1943). New York, NY: Wiley.
- Saxby, D. J., Matthews, G. S., Hitchcock, E. M., & Warm, J. S. (2007). Development of active and passive fatigue manipulations using a driving simulator. *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, 1237-1241. doi: 10.1177/145193120705101839
- Schieben, A., Damböck, D., Kelsch, J., Rausch, H., & Flemisch, F. O. (2008). Haptisches Feedback im Spektrum von Fahrerassistenz und Automation. *Proceedings of the 3rd Conference on Active Safety through Driver Assistance*.
- Schmidt, F. (2010). Detecting and correcting the lies that data tell. *Perspectives on Psychological Science*, 5, 233-242. doi: 10.1177/1745691610369339
- Selye, H. (1973). The evolution of the stress concept. *American Scientist*, 61, 692-699. doi: 10.1016/0002-9149(70)90796-4
- Shariff, S. Z., Bejaimal, S. A. D., Sontrop, J. M., Iansavichus, A. V., Haynes, R. B., Weir, M. A., & Garg, A. X. (2013). Retrieving clinical evidence: A comparison of Pubmed and Google Scholar for quick clinical searches. *Journal of Medical Internet Research*, 15, e164. doi: 10.2196/jmir.2624
- Shen, J., Barbera, J., & Shapiro, C. M. (2006). Distinguishing sleepiness and fatigue: Focus on definition and measurement. *Sleep Medicine Reviews*, 10, 63-76. doi: 10.1016/j.smr.2005.05.004
- Sheridan, T. B. (2004). Driver distraction from a control theory perspective. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46, 587-599. doi: 10.1518/hfes.46.4.587.56807
- Singh, J. (1991). Redundancy in constructs: Problem, assessment, and an illustrative example. *Journal of Business Research*, 22, 255-280. doi: 10.1016/0148-2963(91)90006-J
- Skottke, E-M., Debus, G., Wang, L., & Huestegge, L. (2014). Carryover effects of highly automated convoy driving on subsequent manual driving performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56, 1272-1283. doi: 10.1177/0018720814524594
- Stanley, L. M. (2006). *Haptic and auditory interfaces as a collision avoidance technique during roadway departures and driver perception of these modalities*. (Doctoral dissertation), Montana State University-Bozeman.
- Stanton, N. A., Dunoyer, A., & Leatherland, A. (2011). Detection of new in-path targets by drivers using stop & go adaptive cruise control. *Applied Ergonomics*, 42, 592-601. doi: 10.1016/j.apergo.2010.08.016
- Stanton, N. A., & Marsden, P. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science*, 24, 35-49. doi: 10.1016/S0925-7535(96)00067-7
- Stanton, N. A., Salmon, P. M., Walker, G. H., & Jenkins, D. P. (2010). Is situation awareness all in the mind? *Theoretical Issues in Ergonomics Science*, 11, 29-40. doi: 10.1080/14639220903009938
- Stanton, N. A., Walker, G. H., Young, M. S., Kazi, T. A., & Salmon, P. M. (2007). Changing drivers' minds: The evaluation of an advanced driver coaching system. *Ergonomics*, 50, 1209-1234. doi: 10.1080/00140130701322592
- Stanton, N. A., & Young, M. S. (1998). Vehicle automation and driving performance. *Ergonomics*, 41, 1014-1028. doi: 10.1080/001401398186568

- Stanton, N. A., & Young, M. S. (2000). A proposed psychological model of driving automation. *Theoretical Issues in Ergonomics Science*, 1, 315-331. doi: 10.1080/14639220052399131
- Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics*, 48, 1294-1313. doi: 10.1080/00140130500252990
- Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R. E., . . . Castellan, N. J. (1977). *Tri-level study of the causes of traffic accidents: An overview of final results*. (Report No. DOT HS-805 099), Indiana University, Bloomington, IN.
- Vaa, T. (2007). Modelling driver behaviour on basis of emotions and feelings: Intelligent transport systems and behavioural adaptations. In C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments: Critical issues in driver interactions with intelligent transport systems* (pp. 208-234). London: Springer.
- Vahidi, A., & Eskandarian, A. (2003). Research advances in intelligent collision avoidance and adaptive cruise control. *IEEE Transactions on Intelligent Transportation Systems*, 4, 143-153. doi: 10.1109/TITS.2003.821292
- Van Driel, C. (2007). *Driver support in congestion: An assessment of user needs and impacts on driver and traffic flow*. (Doctoral dissertation), Universiteit Twente.
- Warm, J. S., Matthews, G. S., & Finomore, V. S. (2008). Vigilance, workload and stress. In P. A. Hancock & J. Szalma (Eds.), *Performance under stress* (pp. 225-241). Aldershot: Ashgate Publishing Limited.
- Wheeler, W. A., & Trigs, T. J. (1996). A task analytical view of simulator based training for drivers. *Proceedings of the Road Safety Research and Enforcement Conference 'Effective Partnerships'*, 217-221.
- Wickens, C. D., Lee, J. D., Liu, Y., & Gordon-Becker, S. E. (2004). *An introduction to human factors engineering* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Wilde, G. J. S. (1998). Risk homeostasis theory: An overview. *Injury Prevention*, 4, 89-91. doi: 10.1136/ip.4.2.89
- Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2014). State of science: Mental workload in ergonomics. *Ergonomics*, 58, 1-17. doi: 10.1080/00140139.956151
- Young, M. S., & Stanton, N. A. (1997). Automotive automation: Investigating the impact on drivers' mental workload. *International Journal of Cognitive Ergonomics*, 1, 325-336.
- Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors: The Journal of Human Factors and Ergonomics Society*, 44, 365-375. doi: 10.1518/0018720024497709
- Young, M. S., & Stanton, N. A. (2007). Miles away: Determining the extent of secondary task interference on simulated driving. *Theoretical Issues in Ergonomics Science*, 8, 233-253. doi: 10.1080/14639220601129228
- Young, M. S., Stanton, N. A., & Harris, D. (2007). Driving automation: Learning from aviation about design philosophies. *International Journal of Vehicle Design*, 45, 323-338. doi: 10.1504/IJVD.2007.014908

## Effects of Platooning on Signal-Detection Performance, Workload, Stress, and Fatigue

---

### A Driving Simulator Study

#### Abstract

Platooning, whereby automated vehicles travel closely together in a group, is attractive in terms of safety and efficiency. However, concerns exist about the psychological state of the platooning driver, who is exempted from direct control, yet remains responsible for monitoring the outside environment to detect potential threats. By means of a driving simulator experiment, we investigated the effects on recorded and self-reported measures of workload and stress for three task-instruction conditions: (1) No Task, in which participants had to monitor the road, (2) Voluntary Task, in which participants could do whatever they wanted, and (3) Detection Task, in which participants had to detect red cars. Twenty-two participants performed three 40-min runs in a constant-speed platoon, one condition per run in counterbalanced order. Contrary to some classic literature suggesting that humans are poor monitors, in the Detection Task condition participants attained a high mean detection rate (94.7%) and a low mean false alarm rate (0.8%). Results of the Dundee Stress State Questionnaire indicated that automated platooning was less distressing in the Voluntary Task than in the Detection Task and No Task conditions. In terms of heart rate variability, the Voluntary Task condition yielded a lower power in the low-frequency range relative to the high-frequency range (LF/HF ratio) than the Detection Task condition. Moreover, a strong time-on-task effect was found, whereby the mean heart rate dropped from the first to the third run. In conclusion, participants are able to remain attentive for a prolonged platooning drive, and the type of monitoring task has effects on the driver's psychological state.

Heikooop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2017). Effects of platooning on signal-detection performance: A driving simulator study. *Applied Ergonomics*, 60, 116-127. doi:10.1016/j.apergo.2016.10.016

## Introduction

The concept of a platoon—an actively coordinated, tightly spaced group of vehicles traveling together (Bergenheim et al., 2012; Ren & Green, 1994) has been studied for several decades (e.g., Fenton, Cosgriff, Olson, & Blackwell, 1968; Thorpe, Jochem, & Pomerleau, 1998). Because the vehicles in a platoon are driving with short yet constant headways, substantial benefits are achieved in terms of safety, traffic flow efficiency, and energy consumption (Hochstädter & Cremer, 1997; Karaaslan, Varaiya, & Walrand, 1991; Kunze et al., 2011; Tsugawa, Kato, & Aoki, 2011). Now that sensor, computer, and communication technologies are advancing rapidly, platooning is gaining interest among engineers (e.g., Larson, Liang, & Johansson, 2015; Ploeg, Van de Wouw, & Nijmeijer, 2014) and Human Factors scientists (e.g., Gouy, Wiedemann, Stevens, Brunett, & Reed, 2014; Skottke, Debus, Wang, & Huestegge, 2014).

Platooning often entails both longitudinal and lateral automation (e.g., Bergenheim et al., 2012), and hence no direct inputs by the driver are required. According to current legal frameworks, the driver must always be able to resume manual control (Kim, Perlman, Bogard, & Harrington, 2016; United Nations, 1968). Thus, the role of the driver in a platoon is, at present, ill-defined with, on the one hand, an exemption from control duties and, on the other, the ever-present requirement to be able to reclaim control (see also Norman, 2015). Unless the automated driving technology is legally allowed to drive in all environmental circumstances and is perfectly capable and reliable (or can always bring itself to a minimal-risk condition; Society of Automotive Engineers (SAE), 2014), the possibility remains that the driver has to take over control or modify the automation mode, set-points, or control laws (see also Sheridan, 2011).

Researchers have expressed concerns about the effects of platooning on the driver's psychological state (e.g., Levitan, Golembiewski, & Bloomfield, 1998; Saffarian, De Winter,

& Happee, 2012). Because the driver in a platoon is supervising the automation rather than manually controlling the car, there is the risk of becoming drowsy, mentally underloaded, and fatigued (Cha, 2003; De Waard, Van der Hulst, Hoedemaeker, & Brookhuis, 1999; Saxby, Matthews, Warm, Hitchcock, & Neubauer, 2013; Young & Stanton, 2007). Although automated driving is experienced as effortless, at the same time the drivers are subjected to pressure because they have to remain alert in order to be able to intervene in a critical scenario (Banks, Stanton, & Harvey, 2014; Casner, Hutchins, & Norman, 2016). In fact, the notion that the vehicle is in control but the driver remains responsible for accidents that may occur has been said to be “a formula for extreme stress” (Hancock, 2015p. 138). Furthermore, research has shown that when participants are tasked to monitor a machine in order to detect irregular events, they become frustrated and stressed (Scerbo, 2001; Szalma et al., 2004; Warm, Parasuraman, & Matthews, 2008).

A common advantage within the Human Factors domain is that humans are poor monitors (Hancock & Parasuraman, 1992; Harris, 2002; Kibler, 1965; Pritchett & Lewis, 2010; Sheridan, 1996; Wiener & Curry, 1980), or as Wiener (1985) put it: “After three decades of highly prolific research on human vigilance (Mackie, 1977), we are still making the seemingly contradictory statement: a human being is a poor monitor, but that is what he or she ought to be doing” (p. 87). Farber (1999) pointed out that platooning drivers are unable to remain attentive for prolonged periods and will invariably engage in non-driving tasks. Empirical evidence concurs that drivers of automated vehicles are likely to engage in tasks such as calling on the phone, reading, interacting with a smartphone, or grabbing something from the rear compartment, making them unable to react in time if an emergency happens (Llaneras, Salinger, & Green, 2013; Omae, Hashimoto, Sugamoto, & Shimizu, 2005). It is for this reason that Google removed the steering wheel from their driverless cars (Teller, 2015). However, it is yet unknown whether Google's form of function allocation, in which the human

is engineered out of the control loop, is tenable or legally acceptable (Kim et al., 2016). It certainly runs at odds with how automation has been deployed in complex systems such as aviation, water transport, and process control (see Sheridan, 2002).

Thus far, there appears to be no empirical evidence regarding the psychological state of platooning drivers as a function of monitoring task conditions. Moreover, much of what has been said of humans being poor monitors is based on experiments in which subjects sat in an isolated booth and responded to irregular stimuli having a low signal-to-noise ratio (cf. the highly-cited vigilance experiments by Mackworth, 1948). It is unclear to what extent the results of the classical vigilance paradigm generalize to complex supervisory tasks (Kibler, 1965; Stearman & Durso, 2016). According to a literature review by Cabrall, Happee, and De Winter (2016), there is little overlap between the features of classic vigilance research and published experimental tasks of driving vigilance. A driving simulator study by Funke, Matthews, Warm, and Emo (2007) found that drivers of a semi-automated vehicle actually performed better in a pedestrian-detection task than drivers in a manual control condition. Similarly, an on-road study by Davis, Animashaun, Schoenherr, and McDowell (2008) showed a performance improvement in target-detection performance for automated convoy driving as compared to manual convoy driving.

### **Present Research**

The aim of the present research was to investigate how the monitoring task of drivers in a platoon influences dimensions of stress, workload, and signal-detection performance. Participants were told that a critical situation may occur and that they had to intervene when needed. Three task instructions were compared: (1) 'No Task' (NT), in which no extra task was to be performed, (2) 'Voluntary Task' (VT), in which it was emphasized to the participants that they were free to do whatever they wanted, and (3) 'Detection Task' (DT), in which participants were asked to detect red cars among other traffic in the road environment.

The NT condition assessed the effects of monitoring demands that are similar to those that occur with modern forms of highly automated driving in which drivers should be vigilant for events that the automation cannot handle. The DT condition added extra task demands on top of the baseline monitoring demands, requiring the participant to scan cars in the environment. Conversely, the VT condition created a less demanding situation, allowing the driver to engage in non-driving tasks. The experiment was conducted in a driving simulator, providing a safe and controlled environment in which the traffic behaves identically for all participants.

Based on the aforementioned literature, we expected that the DT condition would yield the highest and the VT condition the lowest scores on stress and workload. In our study, stress dimensions (engagement, distress, & worry) were operationalized with the multi-dimensional Dundee Stress State Questionnaire (DSSQ; Matthews et al., 1999), whereas aspects of workload were assessed with the NASA Task Load Index (TLX). Additionally, we used cardiovascular measures, whereby heart rate was regarded as an indication of stress (Healey & Picard, 2004), and heart rate variability was regarded as an indication of workload (Brookhuis & De Waard, 2010; Cinaz, Arnrich, La Marca, & Tröster, 2013; Fallahi, Motamedzade, Heidarimoghadam, Soltanian, & Miyake, 2016; Jorna, 1992; Luque-Casado, Perales, Cárdenas, & Sanabria, 2016; Suriya-Prakash, John-Preetham, & Sharma, 2015). Moreover, considering the literature about human vigilance performance, we expected that participants in the DT condition would miss a substantial number of red cars. An eye tracker was used to record the percentage of eye-closure as an indicator of task engagement (cf. Körber, Cingel, Zimmermann, & Bengler, 2015; Lal & Craig, 2002; Wierwille, Ellsworth, Wreggit, Fairbanks, & Kirn, 1994).



## Method

### Participants

Twenty-two participants (13 male) aged between 19 and 45 years (M: 29.6; SD: 6.8) with at least 1 year of driving experience (M: 10.0; SD: 6.7) were recruited. Most participants were from the University of Southampton community, with 14 participants being students, researchers, or lecturers at the university, a further four holding an engineering qualification, two being administrators, one being a medicine student, and one a police officer who indicated that driving is part of his profession. In order to retain a typical driving population, we did not apply exclusion criteria regarding personal characteristics that are known to be associated with heart rate variability, such as being a smoker (Barutcu et al., 2005) or general fitness level (Corrales, Torres, Esquivel, Salazar, & Orellana, 2012; Luque-Casado, Zabala, Morales, Mateo-March, & Sanabria, 2013). However, being healthy and having 20/20 vision were inclusion criteria, and given the acute effects of smoking on heart-rate variability (Karakaya et al., 2007; Manzano, Vanderlei, Ramos, & Ramos, 2011), we verified that none of the participants engaged in smoking in between the experimental sessions. Five participants indicated they drove less than once a month, five once a month, three 1-3 days a week, three 4-6 days a week, and six every day in the past 12 months. Seven participants indicated they drove 1-1000 miles, three 1001-5000 miles, six 5001-10,000 miles, four 10,001-20,000 miles, one 20,001-30,000 miles, and one over 50,000 miles in the past 12 months.

All participants in this experiment read and signed a consent form. The study was approved by the Ethics and Research Governance Online of the University of Southampton under submission ID number 13967.

## **Apparatus**

The experiment was performed in the Southampton University Driving Simulator (SUDS; Figure 3.1), a Jaguar XJ Saloon. The simulator ran on STISIM Drive® 3, which is a widely used driving simulator software that allows for custom scenario building (Mets et al., 2011). The simulation was shown on three screens in the front creating a 135-degree field-of-view, one screen at the back for a rear view image, and two side mirror displays.

Seeing Machines FaceLab 5 eye tracker captured the participants' eye movements, and AD Instruments PowerLab26T (consisting of three MLA2505 biopotential electrodes, lead wires with disposable ECG electrode patches, and LabChart 8 software) was used for ECG measurements. 'Normal to Normal' (NN) intervals were extracted by the LabChart 8 software using the standard human ECG mode, with the default 2 standard deviation threshold for detection. Furthermore, a low-pass filter (cut-off frequency 50 Hz) was selected to filter extraneous noise.

## **Environment**

The simulation showed an eight-lane highway (four lanes in either direction) with mild curves and hills. Participants were automatically transported within a five-car platoon, with the third car being the ego car. The environment consisted of sparsely distributed buildings, blank overhead signs, and trees. The starting position of the platoon was on the slow (left) lane of the highway, and the centre-to-centre distance between cars was 9 m (translating into a time headway of about 0.3 s). The longitudinal and lateral movements of all cars of the platoon, including the ego car, were synchronous and fully automated. At the start of each run, the platoon accelerated to 120 km/h and maintained that speed for the entire run. The platoon's lateral movement involved seven overtaking manoeuvres per run by means of a single lane change back and forth.

## **Procedure**

Participants were asked to bring some entertainment they could use while sitting down, like a book or e-reader, something to eat or drink, a laptop, or a mobile phone. These items could be used during one of the conditions of the experiment.

Upon arrival, participants received paper instructions explaining that they would be driving three 40-min runs along a highway in an automated platoon. Furthermore, information on the procedures of the experiment, a consent form, a figure depicting electrode placement, a demographics questionnaire, and the first part of the DSSQ were provided. This pre-task DSSQ queried the participants' current state prior to the simulator runs.

Participants read the instructions and filled out the questionnaires. In addition, the eye tracker was calibrated and the ECG electrodes were attached. The three electrodes were placed in a triangular configuration, with two electrodes placed below the far ends of the collar bones and one electrode over the xiphoid process (males), or one electrode at the top of the sternum and two electrodes below the ribs on both sides (females) (see e.g., Shaffer & Combatalade, 2013).

To indicate readiness to begin the experiment, participants pressed a handheld button, after which the run was started. After each run, the participants received the post-task DSSQ and the TLX. The post-task DSSQ asked the participants about their state while performing the task.

## **Conditions**

The experiment consisted of three 40-min runs, one task condition per run in counterbalanced order. Although no automation malfunction actually occurred, for this is likely to be a rare event in real-life automated platooning, for all three conditions, participants were told that a critical situation may occur and that they had to intervene when needed.

Specifically, before each run, participants received paper instructions, which differed per experimental condition:

- (1) ‘No Task’ (NT), in which participants were asked to monitor the road, to intervene whenever a critical situation appears, and to avoid accidents at all times.
- (2) ‘Voluntary Task’ (VT), in which next to the tasks of the NT condition, participants were told they could do whatever they wanted, including reading, drinking, eating, detecting red cars, or even sleeping, as long as they felt able to intervene at all times.
- (3) ‘Detection Task’ (DT), in which, in addition to the tasks of the NT condition, participants were asked to detect red cars by pressing the handheld button.

Note that the driving task was fully automated, meaning that it was both hands-free and feet-free. Thus, although participants were instructed to intervene if the situation demanded it, they could not intervene at any point during the experiment.



*Figure 3.1.* The Southampton University Driving Simulator (SUDS) during the experiment.

The traffic density in the participants' direction of travel was low, with 23 vehicles distributed across the four lanes that were either overtaking or being overtaken by the platoon during the run. The opposite side of the road contained 5179 vehicles over the duration of the run, comprising a flow of traffic distributed across all four lanes with decreasing density towards the fast lane. The average capacity flow per lane was 1942 cars per hour, which is in line with typical free-flow traffic (Knoop, Hoogendoorn, & Van Zuylen, 2008). Each half run contained either a low rate (LR) or high rate (HR) of red cars, counterbalanced across participants. The red cars could be driving in the participants' own lane (1 being overtaken by the platoon), in the three other lanes (10 overtaking the platoon), as well as in the four lanes among the traffic driving in the opposite direction (59 in total), on predetermined randomly distributed times during the drive. The LR half contained 20 red cars per 20 min, whereas the HR half contained 50 red cars per 20 min, totalling 70 red cars per run. The 70 red cars were conspicuous with respect to the remaining traffic. That is, the colours of the remaining vehicles were never dark orange or brown, but, for example, black, grey, white, or blue.

### **Dependent Measures**

The following dependent measures were calculated per run:

- Hit rate (% of 70 red cars). The hit percentage was calculated automatically by assessing for each of the 70 red cars whether the button was pressed. In the calculation of the hit rate, a two-stage approach was used by distinguishing between the 59 approaching cars and the 11 overtaken/overtaking red cars that could occur simultaneously with the approaching cars. The time that approaching cars were visible was approximately 8 s, including the time they were visible in the mirrors. A hit for an approaching car was defined as a button press between 2 s before until 10 s after the red car could first be seen. This 2-s time buffer before and after appearance of the red

cars was adopted to account for possible synchronisation discrepancies in the data logging. After determining the hits for the 59 approaching cars, it was assessed whether the remaining button presses could be assigned to the overtaken red car (20 s visibility interval) or the 10 overtaking red cars (40 s visibility interval). If a participant pressed a button more than once for the same red car, this counted as a single hit (cf. Mueller & Piper, 2014).

- False alarm rate (% of 53 time intervals). A false alarm was defined as a button press in a time interval in which a red car could not be seen. The maximum possible number of false alarms was 53, and the total time interval in which there was an opportunity for generating a false alarm was 1409 s (corresponding to 58.5% of the eight 300-s intervals).
- Heart rate (bpm), a measure of stress.
- SDNN (ms). This time-domain measure of heart rate variability is defined as the mean of the standard deviation of all NN intervals. A decrease in SDNN is an indication of an increase of workload (e.g., Fallahi et al., 2016; Suriya-Prakash et al., 2015).
- LF/HF ratio (ms). This frequency-domain measure of heart rate variability is defined as the power of the NN interval in the low-frequency (LF) 0.04e0.15 Hz range relative to the high frequency (HF) 0.15e0.40 Hz range. An increase in the LF/HF ratio is an indication of an increase in workload (Cinaz et al., 2013; Suriya-Prakash et al., 2015). Both the SDNN and the LF/HF ratio were calculated from the NN intervals after a default NN artefact filter, using an open-source MATLAB program provided by Vollmer (2015).
- Percentage eyes closed (PERCLOS; %), defined as the percentage of time that the eyes were practically closed across a moving time window of fixed size, excluding blinks.

- DSSQ, a self-report measure of stress states. Version 1.3 (Matthews, Campbell, & Falconer, 2000) was used in this experiment. Standardized change scores for each scale of the DSSQ were calculated as follows:  $(\text{post-score} - \text{pre-score}) / (\text{standard deviation of the pre-score})$  (Helton, Warm, Matthews, Corcoran, & Dember, 2002). The scores for the three DSSQ scales (engagement, distress and worry) were calculated as the means of four subscales (based on Fairclough & Venables, 2005; Heikoop, De Winter, Van Arem, & Stanton, submitted; Matthews, 2014; Matthews et al., 2002). Specifically, engagement consists of the subscales (1) energetic arousal, (2) success motivation, (3) intrinsic motivation, and (4) concentration, distress consists of (5) tense arousal, (6) hedonic tone, (7) control and confidence, and (8) anger/frustration, and worry consists of (9) self-focused attention, (10) self-esteem, (11) task-relevant interference, and (12) task-irrelevant interference. The internal reliability (Cronbach's alpha) of the subscales of the DSSQ ranges from 0.77 to 0.89 (Matthews et al., 2002).
- TLX, as a self-report measure to assess workload. The TLX is the most widely used measure of self-reported workload (De Winter, 2014) and has shown a test-retest reliability of 0.83 across a four-week period (Hart & Staveland, 1988).

The hit rate, false alarm rate, cardiovascular measures, and eye-closure measure were calculated per 5-min interval to be able to assess time-on-task effects within a session, as well as the overall effect of a session. Specifically, each 2413-s long run was divided into eight 300-s segments (10-310 s, 310-610 s, ..., 2110-2410 s; the first 10 s and the last 3 s were discarded). The measures were calculated per segment and subsequently averaged across the eight segments. For the SDNN variable, this approach is equivalent to the SDNN index, a

measure that is robust to low-frequency drifts in the data (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996).

### **Statistical Analyses**

The three conditions were compared with a repeated measures analysis of variance (ANOVA) with a significance level  $\alpha$  of .05. The ANOVA was followed by paired comparisons between the three conditions using a Bonferroni correction. Hence, in the pairwise comparisons, a result was declared statistically significant if the  $p$ -value was smaller than .05/3. Effect sizes between pairs of conditions were calculated as Cohen's  $d_z$  for matched pairs. With a sample size of 22, the power to detect medium ( $d_z = 0.5$ ) and large effects ( $d_z = 0.8$ ) at a significance level of .05/3 was 42% and 86%, respectively (Faul, Erdfelder, Lang, & Buchner, 2007). In order to assess the validity of the workload measures, a Spearman rank-order matrix was constructed between personal characteristics (age, gender, driving experience, driving frequency, mileage) and measures of workload (heart rate variability, self-reported workload).

## **Results**

### **Button Presses, Hits, and False Alarms**

In the VT condition, 11 out of 22 participants took up the detection task, of which 6 performed this task throughout the entire run and 5 partially. In the three runs combined, the button was pressed 2118 times, of which 11 were false alarms, 2073 were hits, and 34 were cases in which a participant pressed the button more than once in a red-car interval. For example, some participants pressed the button twice when a red car was overtaking them: once when the red car appeared in the rear-view mirror and once again when it appeared in the frontal view. Participants never pressed the button twice in a no-red-car interval.



The participants' mean (*SD*) hit rate was 39.9% (42.1%) and 94.7% (4.2%), whereas the mean (*SD*) false alarm rate was 0.2% (0.6%) and 0.8% (1.7%) for the VT and DT conditions, respectively. For the DT condition, the hit rate and false alarm rate translate into a perceptual sensitivity ( $d'$ ) of 4.04 and a response bias ( $\beta$ ) of 5.10 (Stanislaw & Todorov, 1999). In other words, participants were exceptionally well able to discriminate the red cars from the remaining traffic and were highly conservative in responding. Across the 22 participants, the hit rate ranged between 0% and 98.6% for the VT condition, and between 87.1% and 100% for the DT condition. The false alarm rate ranged between 0% and 1.9% for the VT condition, and between 0% and 7.5% for the DT condition.

Figure 3.2 shows the hit rate per 5-min interval. It can be seen that the hit rate remained approximately constant with time, indicating there was no substantial vigilance decrement.

### **Self-Report Questionnaires: DSSQ and TLX**

The results of the DSSQ show that in all three conditions, participants had lost substantial task engagement with respect to the pre-task score (i.e., standardized change scores below zero, see Figure 3.3). Several statistically significant effects between the three conditions were observed for the three DSSQ dimensions. Specifically, the VT condition yielded a lower Distress score than the NT condition. Moreover, NT resulted in higher Worry than VT and DT. The pairwise comparisons of the TLX showed that the DT condition was rated as more mentally demanding than the other two conditions (Figure 3.4). Additionally, participants rated DT as significantly more effortful and frustrating compared to the VT condition.

### Heart Rate

The mean (*SD*) heart rate for the NT, VT, and DT conditions was 70.3 bpm (12.1), 69.4 bpm (10.9), and 70.9 bpm (12.5), respectively. According to a repeated measures ANOVA, the mean heart rate between the three conditions was not significantly different,  $F(2,42) = 0.84, p = .438$ . The pairwise comparisons were not significantly different from each other either (NT vs. VT:  $d_z = 0.16, p = .462$ ; NT vs. DT:  $d_z = -0.12, p = .577$ ; VT vs. DT:  $d_z = -0.26, p = .242$ ).

A follow-up analysis revealed that there was a strong time-on-task effect,  $F(2,42) = 14.1, p < .001$ , with the heart rate dropping significantly from Run 1 ( $M = 72.5, SD = 11.0$ ) to Run 3 ( $M = 67.8, SD = 11.4$ ),  $d_z = 0.96, p < .001$  (see Figure 3.5).

### Heart Rate Variability

The mean (*SD*) SDNN for the NT, VT, and DT conditions was 75.6 ms (29.5), 75.5 ms (30.7), and 75.0 ms (34.9), respectively. The three conditions were not significantly different ( $F(2,42) = 0.01, p = .992$ ), and the pairwise comparisons were not significant either (NT vs. VT:  $d_z = 0.00, p = .985$ ; NT vs. DT:  $d_z = 0.02, p = .910$ ; VT vs. DT:  $d_z = 0.02, p = .921$ ). As with heart rate, a time-on-task effect was found:  $F(2,42) = 11.7, p < .001$ ; the SDNN rose from Run 1 ( $M = 65.6, SD = 27.7$ ) to Run 3 ( $M = 83.1, SD = 33.8$ ),  $d_z = -0.82, p < .001$ .

The mean (*SD*) LF/HF ratio was 1.15 (0.27), 1.07 (0.30), and 1.18 (0.30) for the NT, VT, and DT conditions, respectively. The three conditions were significantly different ( $F(2,42) = 5.78, p = .006$ ), and the pairwise comparisons showed that the VT condition had a lower LF/HF ratio than the DT condition (NT vs. VT:  $d_z = 0.51, p = .025$ ; NT vs. DT:  $d_z = -0.17, p = .429$ ; VT vs. DT:  $d_z = -0.76, p = .002$ ). There was no significant effect of run number on the LF/HF ratio ( $F(2,42) = 1.03, p = .365$ ).

### Percentage Eyes Closed (PERCLOS)

The eye-tracker was often unable to track the participants' eyes. First, we performed a data quality check to examine whether the amount of missing data differed between the three task conditions. For the eye-closure variable, 42% ( $SD = 39\%$ ), 63% ( $SD = 35\%$ ), and 42% ( $SD = 33\%$ ) of data points were missing for the NT, VT, and DT conditions, respectively ( $F(2,42) = 7.09, p = .002$ ; NT vs. VT:  $d_z = -0.59, p = .012$ ; NT vs. DT:  $d_z = -0.02, p = .927$ ; VT vs. DT:  $d_z = 0.73, p = .003$ ). The particularly high amount of missing data for the VT condition may be explained by the fact that participants in this condition were told that they could do whatever they wanted, and therefore strayed from the eye tracker's field-of-view. When selecting the 16 participants who had available eye-closure data for each of their three runs, the mean ( $SD$ ) eye-closure percentages were 5.2% (8.6%), 5.8% (8.9%), and 3.1% (3.7%) for the NT, VT, and DT conditions, respectively,  $F(2,30) = 1.36, p = .271$ .

There were no statistically significant differences between Runs 1, 2, and 3,  $F(2,30) = 0.03, p = .973$ , but a further analysis indicated that eye-closure increased with time-on-task. Specifically, averaged across the three task conditions, the mean eye-closure in the first and second half of the session was 5.1% and 6.9%, respectively. The amount of missing data increased as well, from 47.1% in the first half of the session to 51.0% in the second half of the session.

### Correlation Analysis

Table 3.1 shows the correlation matrix among the personal characteristics, cardiovascular measures, and TLX items. Although interpretation should be done with caution due to the small sample size, several patterns emerge. First, it can be seen that older participants

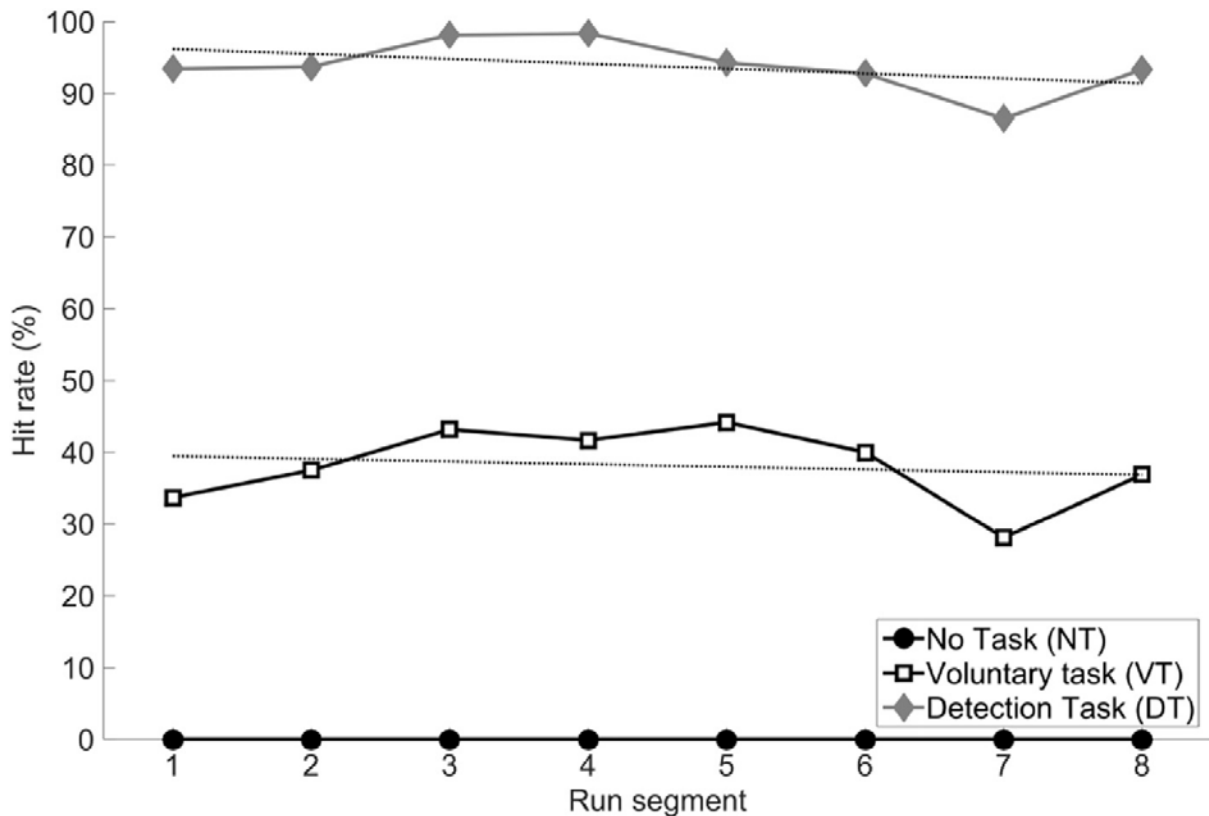


Figure 3.2. Hit rate per 5-min segment during the run, for each of the three experimental conditions. The dotted lines represent linear trend lines. Note: The fluctuation in the hit rate between segments (e.g., low hit rate in Segment 7) is partly attributable to variations in the visibility of the cars (e.g., a red car which was partly/completely covered by another vehicle, or the redness of a car was visible from close by only).

showed indications of lower heart rate variability, which is in line with the literature (Voss, Schroeder, Heitmann, Peters, & Perz, 2015). Second, the heart rate variability measures are not interpretably related to self-reported workload, with one scale of self-reported workload (frustration) in fact showing a statistically significant positive correlation with SDNN. Third, heart rate, SDNN, and the LF/HF ratio are strongly correlated with each other, indicating substantial redundancy of these cardiovascular measures. Fourth, signal detection performance is not significantly related to driving experience and mileage, but it is related to workload, which may be because people who tried harder on the task obtained a better detection performance.

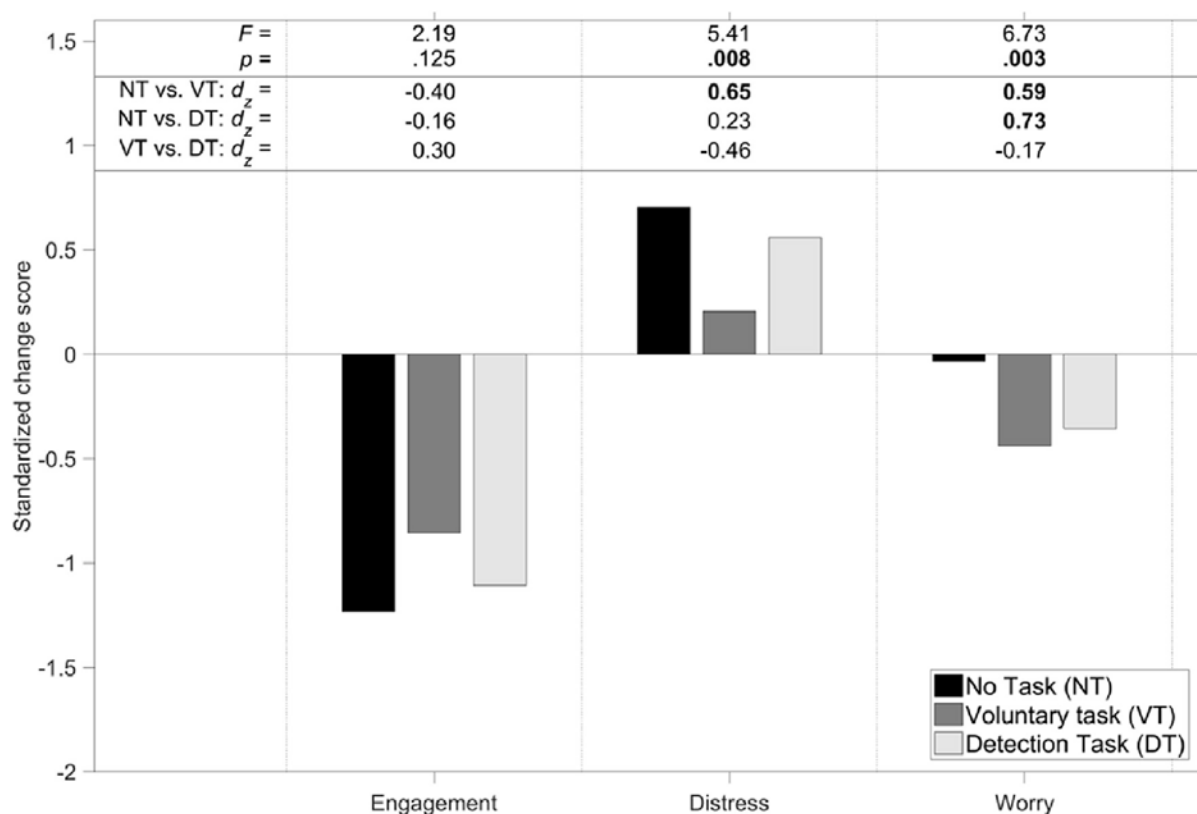


Figure 3.3. Standardized change scores for the DSSQ for the three experimental conditions. Also shown are the  $F$  statistic and  $p$  value for the repeated measures ANOVA and the Cohen's  $d_z$  effect size between the three pairs of conditions. For the ANOVA, the  $p$  value is shown in boldface if  $p < .05$ . For the pairwise comparisons, the  $d_z$  is shown in boldface if  $p < .05/3$ .

## Discussion

The aim of this study was to assess the effect of task instructions on dimensions of workload and stress (engagement, distress, worry) in an automated platoon for extended periods. For each task condition, participants were informed that they had to monitor the road, intervene whenever a critical situation appeared, and avoid accidents at all times. In the Voluntary Task (VT) condition participants were further informed that they could do whatever they wanted. In the No Task (NT) condition, participants had no additional task, and in the Detection Task (DT) condition, participants had the task to detect red cars during the run.

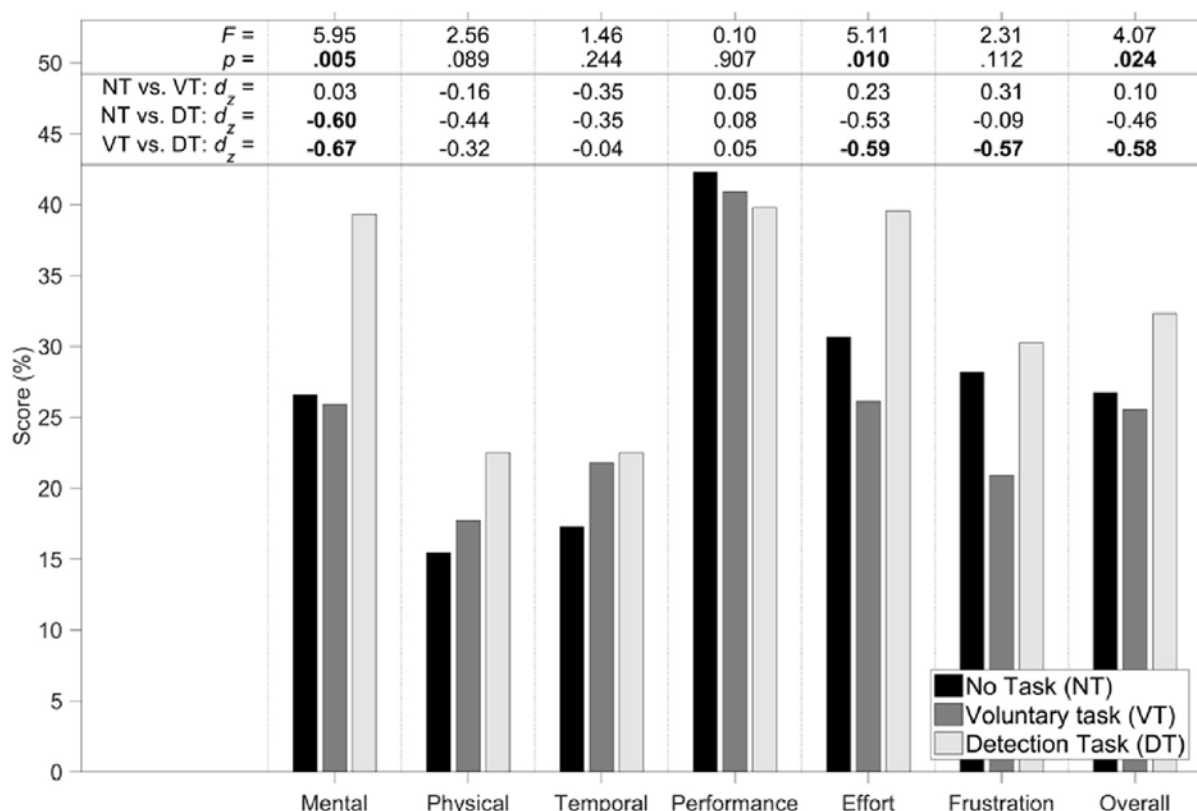


Figure 3.4. Scores on the TLX for the three experimental conditions. The scores are expressed as a percentage and range from Very low (0%) to Very high (100%) for the mental demand, physical demand, temporal demand, effort, and frustration items, and from perfect (0%) to failure (100%) for the performance item. Also shown are the  $F$  statistic and  $p$  value for the repeated measures ANOVA and the Cohen's  $d_z$  effect size between the three pairs of conditions. For the ANOVA, the  $p$  value is shown in boldface if  $p < .05$ . For the pairwise comparisons, the  $d_z$  is shown in boldface if  $p < .05/3$ .

Compared to other automated driving simulator experiments our research is extensive, with three 40-min sessions (2 h of driving in total), giving insights into the longer-term psychological effects of automated platooning. Previous driving simulator research on automated driving involved a total driving time of an average of 1.05 h ( $SD = 0.63$ ) (for an overview see De Winter, Happee, Martens, & Stanton, 2014). In comparison, typical vigilance research, the largest vigilance decrement occurs in the first 15 min (e.g., Mackworth, 1948), and the average trip duration in the U.S. and Europe is 20-30 min (McKenzie & Rapino, 2011; Pasaoglu et al., 2014).

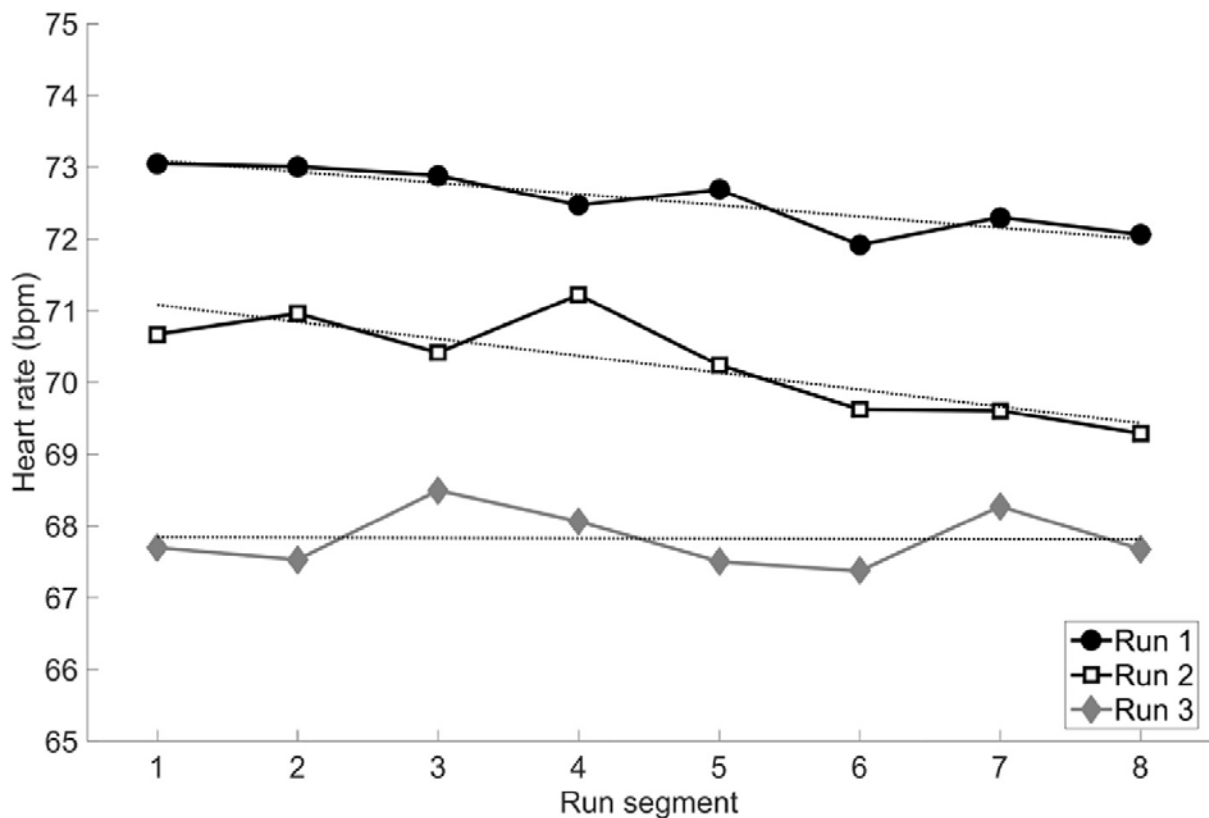


Figure 3.5. Mean heart rate per 5-min segment during the run, for the three runs in chronological order. The dotted lines represent linear trend lines.

### Signal-Detection Performance

Participants adhered to the task instructions: Not one participant pressed the response button in the NT condition, whereas all 22 participants detected at least 61 out of 70 target cars in the DT condition, with a high mean hit rate of 94.7% and a low mean false alarm rate of 0.8%. These results suggest that the saying ‘man is a poor monitor’ does not apply to our automated platooning environment. Although our miss rate was low in comparison to some of the classical vigilance research (e.g., Mackworth, 1948, showing that the miss rate rose from 16% to 28% over a 2-h experiment), missing 5.3% of targets may still pose a high risk in real life if the targets (i.e., the situations that require manual intervention) occur frequently and/or if misses have severe consequences. To illustrate, if each miss of the DT condition resulted in a collision, then there would have been an unacceptably high number of 82 (5.3% of 22 participants x 70 red cars) collisions in 15 h of driving (22 participants x 40 min). The low miss rate, which is in disagreement with our hypothesis, may be caused by the fact that the

Table 3.1

*Spearman rank-order correlations among personal characteristics and workload measures (N = 22, averaged across the three task conditions).*

	Age	Gender	Driving experience	Driving frequency	Mileage	Heart rate	SDNN	LF	HF	LF/HF ratio	TLX – Mental	TLX – Physical	TLX – Temporal	TLX – Performance	TLX – Effort	TLX – Frustration
Age																
Gender	0.00															
Driving experience	0.87	-0.12														
Driving frequency	0.31	-0.23	0.51													
Mileage	0.48	-0.11	0.63	0.79												
Heart rate	0.04	-0.18	0.00	0.35	0.08											
SDNN	-0.32	-0.02	-0.13	-0.14	-0.04	-0.77										
LF	-0.09	-0.04	0.05	0.42	0.22	0.48	-0.15									
HF	-0.32	-0.27	-0.10	-0.06	-0.06	-0.48	0.77	-0.01								
LF/HF ratio	0.21	0.28	0.10	0.30	0.25	0.60	-0.73	0.50	-0.82							
TLX – Mental	-0.15	-0.19	0.10	0.12	-0.01	-0.03	0.02	0.21	0.25	-0.01						
TLX – Physical	0.24	-0.43	0.46	0.24	0.14	-0.14	0.23	0.18	0.35	-0.26	0.53					
TLX – Temporal	-0.17	-0.31	0.05	0.26	0.03	0.04	0.10	0.07	0.27	-0.18	0.69	0.41				
TLX – Performance	-0.25	0.01	-0.22	-0.33	-0.43	0.00	0.06	-0.02	0.01	-0.13	-0.09	0.01	-0.10			
TLX – Effort	-0.02	-0.31	0.24	0.32	0.16	-0.02	0.12	0.38	0.42	-0.11	0.72	0.68	0.48	0.08		
TLX – Frustration	-0.16	0.07	0.17	0.39	0.40	-0.20	0.47	0.27	0.48	-0.25	0.27	0.46	0.33	-0.05	0.39	
Hit rate DT condition	-0.01	-0.24	0.24	0.20	-0.09	0.37	-0.14	0.51	0.05	0.12	0.55	0.46	0.47	0.03	0.31	0.24

*Note. Correlations of magnitude 0.43 or greater are statistically significant from 0 ( $p < 0.05$ )*

environment of the open road is more complex and dynamic than classic vigilance task environments, for which operators may be more likely to remain engaged (for further discussion on task complexity and vigilance see Cummings, Gao, & Thornburg, 2016; Molloy



& Parasuraman, 1996; Montague, Webber, & Adams, 1965). The detection task itself might have been considered engaging as well; Neubauer, Matthews, and Saxby (2014) showed that interacting with media devices is considered to be engaging. Alternatively, it could be that our detection task was easy due to the high saliency of the red cars (i.e., a high signal-to-noise ratio); see Körber et al. (2015) for a similar interpretation regarding the results of their auditory detection task used during partially automated driving.

### **Self-Reported Stress and Workload**

Our results indicate that automated platooning in general (i.e., each of the three task instructions) resulted in a loss of task engagement with respect to the pre-task score (i.e., standardized change scores below zero). These results resemble those of Saxby, Matthews, Hitchcock, and Warm (2007), who found that during a passive (fully automated) condition drivers reported a significant loss of task engagement.

Furthermore, participants found being occupied with a (either voluntary [VT] or mandatory [DT]) task significantly less worrisome than having no task to do (NT). This may be due to the fact that when having no task to do, participants have more opportunity to focus attention on themselves and worry about personal matters. Additionally, the voluntary task yielded lower distress scores than the other two task conditions, which is in agreement with our hypothesis.

Performing the detection task was considered to be somewhat mentally demanding, with a mean TLX mental demand rating of 39% for the DT condition versus 26% for the VT condition, on a scale from very low (0%) to very high (100%) (Figure 3.3). Thus, consistent with our hypothesis, overall, participants found the VT condition the least stressful, and the DT task the most workload-inducing.

By means of personal observations and post-experiment interviews it was found that in the VT condition participants performed actions ranging from eating a sandwich and

interacting with one's mobile phone, to performing the detection task. Indeed, half of participants in the VT condition decided to detect red cars even though this was not required (Figure 3.2). Some of them explained afterwards that they preferred doing the detection task in the VT condition to ensure they remained attentive to the road (and see Miller et al., 2015 for a study showing that engaging in a non-driving task can prevent drowsiness).

### **Heart Rate and Heart Rate Variability**

Our hypotheses regarding workload and stress were not accepted regarding the mean heart rate and SDNN: there were no statistically significant difference between the three conditions. However, heart rate variability in terms of the LF/HF ratio differentiated the VT condition from the other two conditions, supporting the hypothesis that the VT condition was the least workload-inducing. Moreover, we found that the heart rate reduced and SDNN increased substantially during the experiment from Run 1 to Run 3. This suggests that, regardless of the driver's task in an automated platoon, the driver will become less stressed as well as underloaded over time. These results resemble those of Körber et al. (2015), who found significant time-on-task effects for several eye measures such as blink frequency. Collectively, our results indicate that not only the type of task, but also time-on-task has substantial effects on the state of the driver, suggesting that both these effects should be taken into consideration when designing automated driving systems.

### **Limitations**

Certain limitations of this experiment are acknowledged. First, it is impossible to uniquely attribute a specific psychological state (e.g., workload) to a specific overt physiological recording (Cacioppo & Tassinary, 1990). Because psychological states are often substantially correlated (e.g., Desmond & Matthews, 2009), we recognize the risk of 'construct proliferation' (for a review see Heikoop, De Winter, Van Arem, & Stanton, 2016).

The interpretation of heart rate variability is particularly difficult, because the time-domain measure (SDNN) was sensitive to time-on-task (Run 1 vs. Run 3) but not to the type of task, whereas the frequency-domain measure (LF/HF ratio) was sensitive to the type of task (VT vs. DT) but not to time-on-task. It is known that physical activity and posture (e.g., sitting vs. supine or upright position) have substantial effects on heart rate variability (e.g., Bernardi, Valle, Coco, Calciati, & Sleight, 1996; Castiglioni, Parati, Civijian, Quintin, & Di Rienzo, 2009; Pomeranz et al., 1985) and so do individual characteristics such as age (Voss et al., 2015) and gender (Koenig & Thayer, 2016) (see also the correlation analysis in Table 3.1). It cannot be ruled out that these variables have interacted with the experimental conditions, although our within-subject design in which each participant serves as his/her own control ought to be robust against individual differences. The correlation analysis indicated that the heart rate variability measures were not meaningfully related to self-reported workload, but were substantially associated with basic heart rate, raising questions about the discriminant validity of these two measures (see Mehler, Reimer, & Wang, 2011, for a similar observation). The interpretation of heart rate variability as an index of workload or sympathetic versus parasympathetic activity is the topic of ongoing research and debate (e.g., Billman, 2013; Reyes del Paso, Langewitz, Mulder, Roon, & Duschek, 2013; Thayer, Åhs, Fredrikson, Sollers, & Wager, 2012), and we concur with Vollmer (2015) that “the average heart rate and heart rate changes can act as confounding variables” (p. 610). Thus, it appears that heart rate variability is able to discriminate between task conditions, yet the causal pathways and neurophysiological mechanisms remain to be elucidated. This situation is similar to other physiological signals, such as pupil movements, which are known to “empirically reflect variations in central processing load” but for which the physiological causes are indirect and complex (Beatty & Lucero-Wagoner, 2000).

Second, although previous studies have found substantial correlations between driving behaviour in a STISIM simulator and driving behaviour on the road (Bédard, Parkkari, Weaver, Riendeau, & Dahlquist, 2010; Lee, Cameron, & Lee, 2003; Mayhew et al., 2011; Wang et al., 2010), participants' behaviours and physiological states during a simulator experiment are not necessarily representative of real-world platooning. Participants knew that a potential crash would not cause them physical harm and were therefore probably not as stressed as in a real life scenario. Moreover, the fixed-base simulator did not provide vestibular motion feedback, whereas in on-road automated driving, issues of visual-vestibular conflict may influence driver comfort and the uptake of secondary tasks (Diels & Bos, 2016). Thus, a replication of this experiment in a real world-driving scenario is advised.

Third, this research did not contain a control condition in which participants were to drive manually (e.g., Barnard & Lai, 2010; Saxby et al., 2013), nor did we vary the degree of automation reliability and availability (cf. Neubauer, Matthews, Langheim, & Saxby, 2012). For example, it is likely that drivers will become considerably frustrated if automation requires regular manual intervention (De Winter, Stanton, Price, & Mistry, 2016). Conversely, if automation is guaranteed to be safe and no manual intervention is ever to be expected (i.e. Fully Automated Driving or level 4 automation, per BAST and NHTSA definitions, respectively), like traveling in a train, participants are likely to be less stressed.

Fourth, although hazard perception is a critical component of (automated) car driving (Underwood, Crundall, & Chapman, 2011), it is unknown whether detecting red cars is a realistic representation of such task. It remains to be investigated how our results generalize to the anticipation, detection, and response to diverse and realistic hazards that require manual intervention, such as vehicles deviating from their paths, automation malfunction, sudden decelerations of lead cars, or stationary objects.

Fifth, the significant drop of heart rate and rise of heart rate variability might reflect acclimatization to the laboratory setting rather than workload and stress per se. To rule out this effect, multiday sessions of the same participants to the driving simulator are necessary (cf. Beggiato & Krems, 2013; Beggiato, Pereira, Petzoldt, & Krems, 2015; Kazi, Stanton, Walker, & Young, 2007; Pereira, Beggiato, & Petzoldt, 2015).

Sixth, our sample was relatively small ( $N = 22$ ), and therefore the statistical power to detect small effects is low. Specifically, in order to detect effects small effects ( $d_z = 0.2$ ) with 80% power and a significance level of .05/3, 265 participants would be needed (Faul et al., 2007). Only by means of fundamentally different types of research (e.g., large-sample cross-institutional research), it is possible to detect subtle effects of stress and workload that may exist in the population.

Lastly, the eye tracker used in this experiment was prone to missing data, a problem that is common in naturalistic driving tasks (Ahlstrom, Victor, Wege, & Steinmetz, 2012). However, in our case, the amount of missing data itself discriminated more strongly between the task conditions than what was actually measured with the eye-tracker. For future research, we recommend the use of a head-mounted eye-tracker if the goal is to measure eye-closure.

## Conclusion

In conclusion, this research shows that the idea that automated driving puts humans in an extremely stressful monitoring role for which they are “magnificently disqualified” (Hancock, 2015, p. 138) does not generalize to a simulator-based platooning task. Although the VT condition was the least workload-inducing and least stressful, participants in the DT condition remained attentive to the road for 40 min, with an average hit rate of 94.7% on the detection task. Furthermore, our results indicate clear time-on-task effects regarding heart

rate. Similar to Szalma et al. (2004) we recommend that in order to increase performance and reduce stress, both type of task and time-on-task should be considered in system design.

## References

- Ahlstrom, C., Victor, T., Wege, C., & Steinmetz, E. (2012). Processing of eye/head-tracking data in large-scale naturalistic driving data sets. *IEEE Transactions on Intelligent Transportation Systems*, 13, 553–564. doi: 10.1109/TITS.2011.2174786
- Banks, V. A., Stanton, N. A., & Harvey, C. (2014). Sub-systems on the road to vehicle automation: Hands and feet free but not ‘mind’ free driving. *Safety Science*, 62, 505–514. doi: 10.1016/j.ssci.2013.10.014
- Barnard, Y., & Lai, F. (2010). Spotting sheep in Yorkshire: Using eye-tracking for studying situation awareness in a driving simulator. In D. De Waard, A. Axelsson, M. Berglund, B. Peters & C. Weikert (Eds.), *Human factors: A system view of human, technology, and organisation*. (pp. 249–261). Maastricht, The Netherlands: Shaker Publishing.
- Barutcu, I., Esen, A. M., Kaya, D., Turkmen, M., Karakaya, O., Melek, M., . . . Basaran, Y. (2005). Cigarette smoking and heart rate variability: Dynamic influence of parasympathetic and sympathetic maneuvers. *Annals of Noninvasive Electrocardiology*, 10, 324–329. doi: 10.1111/j.1542-474X.2005.00636.x
- Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. Cacioppo, L. Tassinary & G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 142–162). Cambridge, MA: Cambridge University Press.
- Bédard, M., Parkkari, M., Weaver, B., Riendeau, J., & Dahlquist, M. (2010). Assessment of driving performance using a simulator protocol: Validity and reproducibility. *American Journal of Occupational Therapy*, 64, 336–340. doi: 10.5014/ajot.64.2.336
- Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour*, 18, 47–57. doi: 10.1016/j.trf.2012.12.006
- Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F*, 35, 75–84. doi: 10.1016/j.trf.2015.10.005
- Bergenheim, C., Pettersson, H., Coelingh, E., Englund, C., Shladover, S., & Tsugawa, S. (2012). Overview of platooning systems. *Proceedings of the 19th Intelligent Transportation Systems World Congress*.
- Bernardi, L., Valle, F., Coco, M., Calciati, A., & Sleight, P. (1996). Physical activity influences heart rate variability and very-low-frequency components in Holter electrocardiograms. *Cardiovascular Research*, 32, 234–237. doi: 10.1016/0008-6363(96)00081-8
- Billman, G. E. (2013). The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. *Frontiers in Physiology*, 4, 54–58. doi: 10.3389/fphys.2013.00026
- Brookhuis, K. A., & De Waard, D. (2010). Monitoring drivers’ mental workload in driving simulators using physiological measures. *Accident Analysis and Prevention*, 42, 898–903. doi: 10.1016/j.aap.2009.06.001
- Cabral, C. D., Happee, R., & De Winter, J. C. F. (2016). From Mackworth’s clock to the open road: A literature review on driver vigilance task operationalization. *Transportation Research Part F: Traffic Psychology and Behaviour*, 40, 169–189. doi: 10.1016/j.trf.2016.04.001

- Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45, 16–28. doi: 10.1037/0003-066X.45.1.16
- Casner, S. M., Hutchins, E. L., & Norman, D. (2016). The challenges of partially automated driving. *Communications of the ACM*, 59, 70–77. doi: 10.1145/2830565
- Castiglioni, P., Parati, G., Civijian, A., Quintin, L., & Di Rienzo, M. (2009). Local scale exponents of blood pressure and heart rate variability by detrended fluctuation analysis: Effects of posture, exercise, and aging. *IEEE Transactions on Biomedical Engineering*, 56, 675–684. doi: 10.1109/TBME.2008.2005949
- Cha, D. (2003). Driver workload comparison among road sections of automated highway systems. SAE Technical paper (No. 2003-01-0119). doi: 10.4271/2003-01-0119
- Cinaz, B., Arnrich, B., La Marca, R., & Tröster, G. (2013). Monitoring of mental workload levels during an everyday life office-work scenario. *Personal and Ubiquitous Computing*, 17, 229–239. doi: 10.1007/s00779-011-0466-1
- Corrales, M., Torres, B., Esquivel, A., Salazar, M., & Orellana, J. (2012). Normal values of heart rate variability at rest in a young, healthy and active Mexican population. *Health*, 4, 21183. doi: 10.4236/health.2012.47060
- Cummings, M. L., Gao, F., & Thornburg, K. M. (2016). Boredom in the workplace: A new look at an old problem. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 58, 279–300. doi: 10.1177/0018720815609503
- Davis, J., Animashaun, A., Schoenherr, E., & McDowell, K. (2008). Evaluation of semi-autonomous convoy driving. *Journal of Field Robotics*, 25, 880–897. doi: 10.1002/rob.20263
- De Waard, D., Van der Hulst, M., Hoedemaeker, M., & Brookhuis, K. A. (1999). Driver behavior in an emergency situation in the Automated Highway System. *Transportation Human Factors*, 1, 67–82. doi: 10.1207/sthf0101\_7
- De Winter, J. C. F. (2014). Controversy in human factors constructs and the explosive use of the NASA TLX: A measurement perspective. *Cognition, Technology & Work*, 16, 289–297. doi: 10.1007/s10111-014-0275-1
- De Winter, J. C. F., Happee, R., Martens, M., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, 27, 196–217. doi: 10.1016/j.trf.2014.06.016
- De Winter, J. C. F., Stanton, N. A., Price, J. S., & Mistry, H. (2016). The effects of driving with different levels of unreliable automation on self-reported workload and secondary task performance. *International Journal of Vehicle Design*, 70, 297–324. doi: 10.1504/IJVD.2016.076736
- Desmond, P. A., & Matthews, G. S. (2009). Individual differences in stress and fatigue in two field studies of driving. *Transportation Research Part F*, 12, 265–276. doi: 10.1016/j.trf.2008.12.006
- Diels, C., & Bos, J. E. (2016). Self-driving carsickness. *Applied Ergonomics*, 53, 374–382. doi: 10.1016/j.apergo.2015.09.009
- Fairclough, S. H., & Venables, L. (2005). Prediction of subjective states from psychophysiology: A multivariate approach. *Biological Psychology*, 71, 100–110. doi: 10.1016/j.biopsycho.2005.03.007
- Fallahi, M., Motamedzade, M., Heidari Moghadam, R., Soltanian, A. R., & Miyake, S. (2016). Effects of mental workload on physiological and subjective responses during traffic density monitoring: A field study. *Applied Ergonomics*, 52, 95–103. doi: 10.1016/j.apergo.2015.07.009

- Farber, E. I. (1999). Comments on “Driving behaviour in an emergency situation in the Automated Highway System”. *Transportation Human Factors*, 1, 83–85. doi: 10.1207/sthf0101\_8
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. doi: 10.3758/BF03193146
- Fenton, R. E., Cosgriff, R. L., Olson, K. W., & Blackwell, L. M. (1968). One approach to highway automation. *Proceedings of the IEEE Specific Systems and Components*, 56, 556–566. doi: 10.1109/PROC.1968.6344
- Funke, G., Matthews, G. S., Warm, J. S., & Emo, A. K. (2007). Vehicle automation: A remedy for driver stress? *Ergonomics*, 50, 1302–1323. doi: 10.1080/00140130701318830
- Gouy, M., Wiedemann, K., Stevens, A., Brunett, G., & Reed, N. (2014). Driving next to automated vehicle platoons: How do short time headways influence non-platoon drivers’ longitudinal control? *Transportation Research Part F: Traffic Psychology and Behaviour*, 27, 264–273. doi: 10.1016/j.trf.2014.03.003
- Hancock, P. A. (2015). Automobility: The coming use of fully-automated on-road vehicles. *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 137–139. doi: 10.1109/COGSIMA.2015.7108188
- Hancock, P. A., & Parasuraman, R. (1992). Human factors and safety in the design of Intelligent Vehicle-Highway Systems (IVHS). *Journal of Safety Research*, 23, 181–198. doi: 10.1016/0022-4375(92)90001-P
- Harris, D. H. (2002). How to really improve airport security. *Ergonomics in Design: The Quarterly of Human Factors Applications*, 10, 17–22. doi: 10.1177/106480460201000104
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139–183. doi: 10.1016/S0166-4115(08)62386-9
- Healey, J. A., & Picard, R. W. (2004). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6, 156–166. doi: 10.1109/TITS.2005.848368
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2016). Psychological constructs in driving automation: A consensus model and critical comment on construct proliferation. *Theoretical Issues in Ergonomics Science*, 17, 284–303. doi: 10.1080/1463922X.2015.1101507
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (submitted). Acclimatizing to automation: driver workload and stress during partially automated car following in real traffic.
- Helton, W. S., Warm, J. S., Matthews, G. S., Corcoran, K. J., & Dember, W. N. (2002). Further tests of an abbreviated vigilance task: Effect of signal salience and jet aircraft noise on performance and stress. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46, 1546–1550. doi: 10.1177/154193120204601704
- Hochstädter, A., & Cremer, M. (1997). Investigating the potential of convoy driving for congestion dispersion. *Proceedings of the IEEE Conference on Intelligent Transportation Systems*, 735–740. doi: 10.1109/ITSC.1997.660565
- Jorna, P. G. A. M. (1992). Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. *Biological Psychology*, 34, 237–257. doi: 10.1016/0301-0511(92)90017-O



- Karaaslan, U., Varaiya, P., & Walrand, J. (1991). Two proposals to improve freeway traffic flow. *Proceedings of the American Control Conference* (pp. 2539-2544). Boston, MA.
- Karakaya, O., Barutcu, I., Kaya, D., Esen, A.M., Saglam, M., Melek, M., & Kaymaz, C. (2007). Acute effect of cigarette smoking on heart rate variability. *Angiology*, 58, 620-624. doi: 10.1177/0003319706294555
- Kazi, T. A., Stanton, N. A., Walker, G. H., & Young, M. S. (2007). Designer driving: Drivers' conceptual models and level of trust in adaptive cruise control. *International Journal of Vehicle Design*, 45, 339-360. doi: 10.1504/IJVD.2007.014909
- Kibler, A. W. (1965). The relevance of vigilance research to aerospace monitoring tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 7, 93-99. doi: 10.1177/001872086500700202
- Kim, A., Perlman, D., Bogard, D., & Harrington, R. (2016). *Review of federal motor vehicle safety standards (FMVSS) for automated vehicles: Identifying potential barriers and challenges for the certification of automated vehicles using existing FMVSS*. (Report No. DOT-VNTSC-OSTR-16-03), National Highway Traffic Safety Administration, Washington, DC.
- Knoop, V., Hoogendoorn, S., & Van Zuylen, H. (2008). Capacity reduction at incidents: empirical data collected from a helicopter. *Transportation Research Record: Journal of the Transportation Research Board*, 2071, 19-25. doi: 10.3141/2071-03
- Koenig, J., & Thayer, J. F. (2016). Sex differences in healthy human heart rate variability: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, 64, 288-310. doi: 10.1016/j.neubiorev.2016.03.007
- Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics*, 365-371. doi: 10.1016/j.promfg.2015.07.499
- Kunze, R., Haberstroh, M. A., Hauck, E., Ramakers, R., Henning, K., & Jeschke, S. (2011). Automated truck platoons on motorways – A contribution to the safety on roads. In S. Jeschke, I. Isenhardt & K. Henning (Eds.), *Automation, Communication and Cybernetics in Science and Engineering 2009/2010*: Springer: Berlin, Heidelberg.
- Lal, S. K. L., & Craig, A. (2002). Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiology*, 39, 313-321. doi: 10.1017/S004857720139095
- Larson, J., Liang, K. Y., & Johansson, K. H. (2015). A distributed framework for coordinated heavy-duty vehicle platooning. *IEEE Transactions on Intelligent Transportation Systems*, 16, 419-429. doi: 10.1109/TITS.2014.2320133
- Lee, H. C., Cameron, D., & Lee, A. H. (2003). Assessing the driving performance of older adult drivers: On-road versus simulated driving. *Accident Analysis and Prevention*, 35, 797-803. doi: 10.1016/S0001-4575(02)00083-0
- Leviton, L., Golembiewski, G., & Bloomfield, J. R. (1998). Human Factors issues for Automated Highway Systems. *ITS Journal – Intelligent Transportation Systems Journal: Technology, Planning, and Operations*, 4, 21-47. doi: 10.1080/10248079808903735
- Llaneras, R. E., Salinger, J., & Green, C. A. (2013). Human Factors issues associated with limited ability autonomous driving systems: Drivers' allocation of visual attention to the forward road. *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (pp. 92-98). Bolton Landing, NY.
- Luque-Casado, A., Perales, J. C., Cárdenas, D., & Sanabria, D. (2016). Heart rate variability and cognitive processing: The autonomic response to task demands. *Biological Psychology*, 113, 83-90. doi: 10.1016/j.biopsycho.2015.11.013

- Luque-Casado, A., Zabala, M., Morales, E., Mateo-March, M., & Sanabria, D. (2013). Cognitive performance and heart rate variability: The influence of fitness level. *PLoS One*, 8, e56935. doi: 10.1371/journal.pone.0056935
- Mackie, R. R. (1977). *Vigilance: Theory, operational performance, and physiological correlates*. New York, NY: Plenum.
- Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1, 6-21. doi: 10.1080/17470214808416738
- Manzano, B. M., Vanderlei, L. C. M., Ramos, E. M., & Ramos, D. (2011). Acute effects of smoking on autonomic modulation: Analysis by Poincaré plot. *Arquivos Brasileiros de Cardiologia*, 96, 154-160. doi: 10.1590/S0066-782X2011005000013
- Matthews, G. S. (2014). *Task-induced fatigue, operator engagement and the control of attention*. PowerPoint slides.
- Matthews, G. S., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., . . . Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, 2, 315-340. doi: 10.1037/1528-3542.2.4.315
- Matthews, G. S., Campbell, S.E., & Falconer, S. (2000). Personality, energetics and individual differences in motivation. *Proceedings of the Tenth European Conference on Personality*. Krakow, Poland.
- Matthews, G. S., Joyner, L., Gilliland, K., Campbell, S., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state 'Big Three'? In I. Mervielde, I. J. Deary, F. De Fruyt & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 335-350). Tilburg: Tilburg University Press.
- Mayhew, D. R., Simpson, H. M., Wood, K. M., Lonero, L., Clinton, K. M., & Johnson, A. G. (2011). On-road and simulated driving: Concurrent and discriminant validation. *Journal of Safety Research*, 42, 267-275. doi: 10.1016/j.jsr.2011.06.004
- McKenzie, B., & Rapino, M. (2011). Commuting in the United States: 2009. *US Department of Commerce, Economics and Statistics Administration, US Census Bureau*.
- Mehler, B., Reimer, B., & Wang, Y. (2011). A comparison of heart rate and heart rate variability indices in distinguishing single-task driving and driving under secondary cognitive workload. *Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 590-597). Lake Tahoe, CA.
- Mets, M. A., Kuipers, E., Senerpont Domis, L. M., Leenders, M., Olivier, B., & Verster, J. C. (2011). Effects of alcohol on highway driving in the STISIM driving simulator. *Human Psychopharmacology: Clinical and Experimental*, 26, 434-439. doi: 10.1002/hup.1226
- Miller, D., Sun, A., Johns, M., Ive, H., Sirkin, D., Aich, S., & Ju, W. (2015). Distraction becomes engagement in automated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 59, 1676-1680. doi: 10.1177/1541931215591362
- Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 38, 311-322. doi: 10.1177/001872089606380211
- Montague, W. E., Webber, C. E., & Adams, J. A. (1965). The effects of signal and response complexity on eighteen hours of visual monitoring. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 7, 163-172. doi: 10.1177/001872086500700208

- Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, 222, 250-259. doi: 10.1016/j.jneumeth.2013.10.024
- Neubauer, C., Matthews, G. S., Langheim, L., & Saxby, D. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 54, 734-746. doi: 10.1177/0018720811423261
- Neubauer, C., Matthews, G. S., & Saxby, D. (2014). Fatigue in the automated vehicle: Do games and conversation distract or energize the driver? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58, 2053-2057. doi: 10.1177/1541931214581432
- Norman, D. A. (2015). The human side of automation. In G. Meyer & S. Beiker (Eds.), *Road vehicle automation 2* (pp. 73-79): Springer International Publishing Switzerland.
- Omae, M., Hashimoto, N., Sugamoto, T., & Shimizu, H. (2005). Measurement of driver's reaction time to failure of steering controller during automatic driving. *Review of Automotive Engineering*, 26, 213-215.
- Pasaoglu, G., Fiorello, D., Martino, A., Zani, L., Zubaryeva, A., & Thiel, C. (2014). Travel patterns and the potential use of electric cars—Results from a direct survey in six European countries. *Technological Forecasting and Social Change*, 87, 51–59. doi: 10.1016/j.techfore.2013.10.018
- Pereira, M., Beggiato, M., & Petzoldt, T. (2015). Use of adaptive cruise control functions on motorways and urban roads: Changes over time in an on-road study. *Applied Ergonomics*, 50, 105-112. doi: 10.1016/j.apergo.2015.03.002
- Ploeg, J., Van de Wouw, N., & Nijmeijer, H. (2014). Lp string stability of cascaded systems: Application to vehicle platooning. *Proceedings of the IEEE Transactions on Control Systems Technology*, 22, 786-793. doi: 10.1109/TCST.2013.2258346
- Pomeranz, B., Macaulay, R. J. B., Caudill, M. A., Kutz, I., Adam, D., Gordon, D., . . . Benson, H. (1985). Assessment of autonomic function in humans by heart rate spectral analysis. *American Journal of Physiology-Heart and Circulatory Physiology*, 248, H151–H153.
- Pritchett, A. R., & Lewis, D. S. (2010). Cognitive engineering and aviation safety. *Encyclopedia of Aerospace Engineering*. doi: 10.1002/9780470686652.eae508
- Ren, W., & Green, D. (1994). Continuous platooning: A new evolutionary and operating concept for automated highway systems. *Proceedings of the American Control Conference*, 11(20), 21-25. doi: 10.1109/ACC.1994.751685
- Reyes del Paso, G. A., Langewitz, W., Mulder, L. J., Roon, A., & Duschek, S. (2013). The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: A review with emphasis on a reanalysis of previous studies. *Psychophysiology*, 50, 477-487. doi: 10.1111/psyp.12027
- Saffarian, M., De Winter, J. C. F., & Happee, R. (2012). Automated driving: Human-factors issues and design solutions. *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting*, 2296-2300. doi: 10.1177/1071181312561483
- Saxby, D. J., Matthews, G. S., Hitchcock, E. M., & Warm, J. S. (2007). Development of active and passive fatigue manipulations using a driving simulator. *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, 1237-1241. doi: 10.1177/145193120705101839
- Saxby, D. J., Matthews, G. S., 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. *Journal of Experimental Psychology: Applied*, 19, 287-300. doi: 10.1037/a0034386

- Scerbo, M. W. (2001). Stress, workload, and boredom in vigilance: A problem and an answer. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 267-278). Mahwah, NJ: Lawrence Erlbaum Associates.
- Shaffer, F., & Combatalade, D. C. (2013). Don't add or miss a beat: A guide to cleaner Heart Rate Variability recordings. *Biofeedback*, 41, 121-130. doi: 10.5298/1081-5937-41.3.04
- Sheridan, T. B. (1996). Allocating functions among humans and machines. In D. Beevis, P. Essens & H. Schuffel (Eds.), *Improving function allocation for integrated systems design* (pp. 7-27). (Report No. CSERIAC SOAR 96-01), Wright Patterson AFB: Crew Systems Ergonomics Information Analysis Center.
- Sheridan, T. B. (2002). *Humans and automation: Systems design and research issues*. New York, NY: John Wiley & Sons, Inc.
- Sheridan, T. B. (2011). Adaptive automation, level of automation, allocation authority, supervisory control, and adaptive control: Distinctions and modes of adaptation. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 41, 662-667. doi: 10.1109/TSMCA.2010.2093888
- Skottke, E-M., Debus, G., Wang, L., & Huestegge, L. (2014). Carryover effects of highly automated convoy driving on subsequent manual driving performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56, 1272-1283. doi: 10.1177/0018720814524594
- Society of Automotive Engineers (SAE). (2014). Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (Standard No. J3016) SAE International.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31, 137-149. doi: 10.3758/BF03207704
- Stearman, E. J., & Durso, F. T. (2016). Vigilance in a dynamic environment. *Journal of Experimental Psychology: Applied*, 22, 107-123. doi: 10.1037/xap0000075
- Suriya-Prakash, M., John-Preetham, G., & Sharma, R. (2015). Is heart rate variability related to cognitive performance in visuospatial working memory? *PeerJ PrePrints*, 3, e1711. doi: 10.7287/peerj.preprints.1377v1
- Szalma, J. L., Warm, J. S., Matthews, G. S., Dember, W. N., Weiler, E. M., & Meier, A. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46, 219-233. doi: 10.1518/hfes.46.2.219.37334
- Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: Standards of measurement, physiological interpretation and clinical use. *Circulation*, 93, 1043-1065. doi: 10.1161/01.CIR.93.5.1043
- Teller, A. (2015). *Moonshots and reality*. Keynote Presentation given at South by Southwest (SXSW), Austin, TX, March 17.
- Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews*, 36, 747-756. doi: 10.1016/j.neubiorev.2011.11.009
- Thorpe, C., Jochem, T., & Pomerleau, D. (1998). The 1997 automated highway free agent demonstration. *Proceedings of the IEEE Conference on Intelligent Transportation Systems*, 496-501. doi: 10.1109/ITSC.1997.660524

- Tsugawa, S., Kato, S., & Aoki, K. (2011). An automated truck platoon for energy saving. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4109-4114. doi: 10.1109/IROS.2011.6094549
- Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14, 435-446. doi: 10.1016/j.trf.2011.04.008
- United Nations. (1968). *Convention on road traffic*. Vienna, Austria.
- Vollmer, M. (2015). A robust, simple and reliable measure of heart rate variability using relative RR intervals. *IEEE Conference on Computing in Cardiology Conference (CinC)*, 609-612. doi: 10.1109/CIC.2015.7410984
- Voss, A., Schroeder, R., Heitmann, A., Peters, A., & Perz, S. (2015). Short-term heart rate variability - Influence of gender and age in healthy subjects. *PLOS ONE*, 10, e0118308. doi: 10.1371/journal.pone.0118308
- Wang, Y., Mehler, B., Reimer, B., Lammers, V., D'Ambrosio, L. A., & Coughlin, J. F. (2010). The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing. *Ergonomics*, 53, 404-420. doi: 10.1080/00140130903464358
- Warm, J. S., Parasuraman, R., & Matthews, G. S. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, 50, 433-441. doi: 10.1518/001872008X312152
- Wiener, E. L. (1985). Beyond the sterile cockpit. *Human Factors*, 27, 75-90. doi: 10.1177/001872088502700107
- Wiener, E. L., & Curry, R. E. (1980). Flight-deck automation: Promises and problems. *Ergonomics*, 23, 995-1011. doi: 10.1080/00140138008924809
- Wierwille, W. W., Ellsworth, L. A., Wreggit, S. S., Fairbanks, R. J., & Kirn, C. L. (1994). *Research on vehicle-based driver status/performance monitoring: Development, validation, and refinement of algorithms for detection of driver drowsiness*. (Report No. DOT HS 808 247), National Highway Traffic Safety Administration, Washington, DC.
- Young, M. S., & Stanton, N. A. (2007). What's skill got to do with it? Vehicle automation and driver mental workload. *Ergonomics*, 50, 1324-1339. doi: 10.1080/00140130701318855

# Effects of Mental Demands on Situation Awareness during Automated Platooning

---

## A Driving Simulator Study

### Abstract

Previous research shows that drivers of automated vehicles are likely to engage in visually demanding tasks, causing impaired situation awareness. How mental task demands affect situation awareness is less clear. In a driving simulator experiment, 33 participants completed three 40-minute runs in an automated platoon, each run with a different level of mental task demands. Results showed that high task demands (i.e., performing a 2-back task, a working memory task in which participants had to recall a letter, presented two letters ago) induced high self-reported mental demands (71% on the NASA Task Load Index), while participants reported low levels of self-reported task engagement (measured with the Dundee Stress State Questionnaire) in all three task conditions in comparison to the pre-task measurement. Participants' situation awareness, as measured using a think-out-loud protocol, was affected by mental task demands, with participants being more involved with the mental task itself (i.e., to remember letters) and less likely to comment on situational features (e.g., car, looking, overtaking) when task demands increased. Furthermore, our results shed light on temporal effects, with heart rate decreasing and self-constructed mental models of automation growing in complexity, with run number. It is concluded that mental task demands reduce situation awareness, and that not only type-of-task, but also time-on-task, should be considered in Human Factors research of automated driving.

Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2017). *Effects of mental demands on situation awareness during automated platooning: A driving simulator study*. Manuscript submitted for publication.

## **Introduction**

### **Emergence of Automated Platooning**

Automated vehicles are being developed at a rapid pace, and systems are emerging that automate longitudinal and lateral control simultaneously. A specific concept that combines longitudinal and lateral automated driving is a platoon, a group of vehicles that drive closely together in a coordinated automated manner (Bergenheim et al., 2012; Ren & Green, 1994). Platooning of automated vehicles offers advantages compared to manual driving in terms of safety, road capacity, and fuel economy (e.g., Axelsson, 2017; Kunze et al., 2011; Larson, Liang, & Johansson, 2015).

### **The Task of Drivers in a Platoon**

Because a platoon may involve time headways as small as 0.3 s (Ploeg, Van de Wouw, & Nijmeijer, 2014), which is at the limit of human reaction time capabilities, it is unreasonable to expect that platooning drivers take over control safely in case of emergency. Nonetheless, the possibility does exist that platooning drivers have to take over control, for example in case of hardware failure (De Waard, Van der Hulst, Hoedemaeker, & Brookhuis, 1999) or in case of a voluntary driver-initiated transition (e.g., exiting or leaving the platoon; see Levitan, Golembiewski, & Bloomfield, 1998; Nilsson, 2014). Accordingly, the question arises what happens to drivers' psychological readiness after having driven in a platoon for some time.

### **Previous Research on Situation Awareness and Automated Driving**

Previous driving simulator research has found that drivers of an automated car experience low levels of workload when having nothing to do (Cha, 2003; De Waard et al., 1999; Heikoop, De Winter, Van Arem, & Stanton, 2017; Young & Stanton, 2007). Heikoop et al. (2017) found that participants in a platoon were still able to remain attentive and detect the majority (95%) of irregularly occurring stimuli (red cars) during 40 minutes of driving, if

tasked to do so (Heikoop et al., 2017). This indicated that participants are able to retain situation awareness despite low workload.

However, when participants were allowed to engage in secondary tasks, only about 40% of the targets were detected (Heikoop et al., 2017); many participants engaged in visually demanding tasks such as eating their lunch or using their phone, rather than to attend to the roadway. Other research has also found that drivers of highly automated cars are likely to pick up visual tasks such as texting, reading, and watching a DVD (Llaneras, Salinger, & Green, 2013; Omae, Hashimoto, Sugamoto, & Shimizu, 2005), as well as mentally demanding tasks such as calling on a phone, or listening to the radio (Carsten, Lai, Barnard, Jamson, & Merat, 2012; Kyriakidis, Happee, & De Winter, 2015). Although it is clear that visual demands impair situation awareness, it is less clear to what extent mental task demands (i.e., engaging in a mentally demanding secondary task) influence situation awareness in automated driving.

Previous research (Gold, Berisha, & Bengler, 2015; Louw, Madigan, Carsten, & Merat, 2016; Petermeijer, Cieler, & De Winter, 2017) has found that a visually demanding task (e.g., performing a SuRT task, or looking at a video) has a stronger negative effect on drivers' take-over performance than a mentally demanding task (i.e., performing an n-back task). Several driving simulator studies have even found that mental demands induced by a verbal task yielded *improved* steering behaviour and lane keeping performance (Atchley, Chan, & Gregersen, 2014; Saxby, Matthews, & Neubauer, 2017; Verwey & Zaidel, 1999). These findings can be explained with the malleable resource theory (MART; Young & Stanton, 2002) as the added demands of using a cell phone could expand the available resource pools. How a mentally demanding secondary task influences driving performance may depend on its frequency and duration of use (Neubauer, Matthews, & Saxby, 2014), the relevance of its contents to the driving task (Saxby et al., 2017), and whether the secondary



task is at all engaging (Bueno et al., 2016). Although mental demands in certain cases may improve driving performance and reaction times, it remains to be clarified whether mental demands are not harmful for higher levels of situation awareness. Indeed, previous research in manual driving suggests that mental demands (i.e., listening to auditory instructions from a navigation system) reduce level 2 (comprehension) and level 3 (anticipation) situation awareness, whereas visual demands (i.e., identifying a target symbol on a tablet display every 10 seconds) impair all three levels of situation awareness (Rogers, Zhang, Kaber, & Liang, 2011). Similarly, in an adaptive cruise control (ACC) study, the cognitive task of using the cell phone showed deleterious effects on drivers' level 3 situation awareness (Ma & Kaber, 2005).

### **Aim of this Research**

The aim of the present research was to investigate the impact of a mental secondary task on driver situation awareness during platooning. Participants performed three 40-minute platooning runs in a simulator, and their situation awareness, and self-reported levels of workload and associated physiological states were measured. We hypothesized that mental secondary task demands, as induced by a verbal N-back task, would have a negative effect on participants' situation awareness. We also probed participants' mental models (i.e., the participants' understanding of the working mechanisms) of the automation after each run. A mental model is an important concept that develops with driving experience (Beggiato & Krems, 2013) and which is considered to be a facilitator of situation awareness (e.g., Biester, 2008; Endsley, 1995; Sarter & Woods, 1991; see Heikoop, De Winter, Van Arem, & Stanton, 2016 for a review). Because car manuals are hardly read (Mehlenbacher, Wogalter, & Laughery, 2002), it appears realistic to investigate drivers' situation awareness and mental models without informing participants about the workings of the automated system and the environmental cues of relevance. Thus, in contrast to most other research using normative

approaches by comparing to a ground truth (e.g., Situation Awareness Global Assessment Technique; Endsley, 1988), we used concurrent think-aloud protocols (Ericsson & Simon, 1980; Salmon, Lenne, Walker, Stanton, & Filtness, 2014) and self-reported concept maps (Revell & Stanton, 2012).

## Methods

### Participants

Thirty-three participants (19 male, 14 female) aged between 18 and 66 years ( $M = 31.0$ ;  $SD = 13.0$ ) with at least 1 year of driving experience ( $M = 12.5$ ;  $SD = 13.1$ ) participated in this experiment. All participants were recruited from the University of Southampton campus through an advertisement on the university internal webpage. Inclusion criteria for participants to partake in this experiment were that they had to hold a full driver's license, be native English speakers, have normal vision and good hearing, and be in a healthy condition. Participants received a monetary incentive of £20.

Of the participants who took part, 14 indicated to be students and/or researchers, 4 to be in a managerial position, 4 in a supporting or advisory position, 2 to be administrators, and 4 to have other types of professions. The remaining 5 participants had no profession or did not disclose one. Eleven participants indicated to drive daily, 7 participants reported 4–6 days a week, 6 reported 1–3 days a week, 5 reported once a month, 2 reported less than once a month, and 2 reported they never drove in the past 12 months. Those 2 also indicated to have 0 mileage over the last 12 months, while 8 drove 1–1,000 miles, 6 drove 1,001–5,000 miles, 12 drove 5,001–10,000 miles, and 5 drove 10,001–20,000 miles. No-one indicated to have driven more than 20,000 miles in the past 12 months.

The study was approved by the Ethics and Research Governance Online of the University of Southampton under submission ID number 18070, and all participants provided written informed consent.

### **Apparatus**

The simulator and electrocardiography (ECG) equipment used for this experiment were identical to a previous study by Heikoop et al. (2017). The experiment was conducted in the Southampton University Driving Simulator (SUDS). The simulator consisted of a Jaguar XJ Saloon and ran on STISIM Drive 3 software. The simulation was presented on three front screens creating a 135-degree field-of-view, one back screen for a rear view image, and two side mirror displays.

The ECG measurements were performed with AD Instruments PowerLab26T, three MLA2505 biopotential electrodes, and LabChart 8 software. ‘Normal to Normal’ (NN) intervals were extracted by the LabChart 8 software using the standard human ECG mode. Ergoneers’ Dikablis Professional head-mounted eye tracker with D-Lab software was used to capture eye movements.

### **Environment**

The experiment entailed the same virtual environment as Heikoop et al. (2017). Specifically, the environment consisted of an eight-lane highway (four lanes in either direction) with mild curves and hills. Participants were transported automatically in a five-car platoon, with the third car being the participant’s car. The time headway between cars was approximately 0.3 s. The longitudinal and lateral movements of all cars of the platoon, including the participant’s car, were identical and fully automated. At the start of each run, the platoon accelerated to 120 km/h and maintained this speed for the entire run. The platoon

made seven overtaking manoeuvres per run by means of a single lane change to the adjacent lane, and back.

### **Procedure**

Upon arrival, participants received paper instructions explaining that they would be driving three 40-min runs on a highway in an automated platoon. Furthermore, information on the procedures of the experiment, condition-specific instructions (see Section 2.5), a consent form, a figure depicting electrode placement, a demographics questionnaire, and the short pre-task version of the Dundee Stress State Questionnaire (DSSQ; Matthews, Emo, & Funke, 2005) were provided. This pre-task DSSQ queried the participants' current stress state, whereas the post-task DSSQs queried the participants' stress state regarding the task they were performing in the preceding session.

Participants read the instructions and completed the questionnaires. In addition, the ECG electrodes were attached. The three electrodes were placed in a triangular configuration, with two electrodes placed below the far ends of the collar bones and one electrode over the xiphoid process (males), or one electrode at the top of the sternum and two electrodes below the ribs on both sides (females) (see e.g., Shaffer & Combatalade, 2013).

Once the forms were completed, the Quick Association Check (QuACk) was administered for measuring the participants' mental model of automated driving (Revell & Stanton, 2016). The QuACk method consists of three steps, namely (1) asking the participants about their prior experience with the technology, (2) asking them about their common use of said technology, and (3) asking them to create a pen-and-paper mental model of how they think the technology works. In order to conduct the experiment within a reasonable time frame, we applied only step 3 of the QuACk. Specifically, participants were provided with an A3 sheet of paper, a pen, and Post-It notes. They were instructed to create a concept map of how they think automated driving works by writing down concepts they thought were present

in an automated driving system on the Post-It notes, placing the Post-It notes on the A3 sheet, and indicating with arrows drawn on the A3 sheet how they think these concepts link to each other. To minimize bias in the data collection of the mental models, participants were not assisted in creating ideas for concepts or links (Revell & Stanton, 2012). Furthermore, it was emphasized that there is no wrong or right answer.

As a final step in the preparation, participants were asked to wear the head-mounted eye tracker, after which it was calibrated. To indicate readiness to begin the experiment, participants pressed a handheld button, after which the first out of three runs was started. After each run, the participants received the post-task DSSQ, the NASA Task Load Index (TLX), and instructions for the next run. Once the questionnaires were completed, participants received the QuACk map back and were asked whether they want to add, remove or alter something based on the experience they had gained during the preceding run.

### **Conditions**

The experiment consisted of three 40-min runs, one task condition per run in counterbalanced order. Prior to each run, participants were told that they had to monitor the road and intervene when a critical situation appeared. Furthermore, they were required to “think out loud” in 2-min intervals (i.e., 2 min of speaking followed by 2 min of silence, etc.), meaning that they had to say out loud whatever they were thinking of at that moment, regardless of its content. This resulted in ten 2-min think-out-loud periods per run, which were used to assess participants’ situation awareness. Participants were alerted of the start and end of a 2-min think-out-loud interval by means of a pre-recorded voice saying “please resume protocol” and “please stop protocol”.

Before each run, participants received paper instructions which differed per experimental condition:

- 1) 'Low Task Demands' (LTD), in which no additional tasks were provided other than those mentioned above.
- 2) 'Medium Task Demands' (MTD), in which participants were, next to the tasks in the LTD condition, *encouraged*, but not *required* to perform a 2-back task by repeating the consonant that was uttered 2 letters ago. The interval between two consonants was exactly 15 seconds and continued throughout the entire run.
- 3) 'High Task Demands' (HTD), in which the participants were *required* to, next to all the basic tasks of the LTD condition, perform the 2-back task as in the MTD condition.

Note that, despite the fact that the participants were told to intervene when required, no intervention was possible throughout the experiment.

### **Dependent Measures**

The following dependent measures were calculated per run:

- DSSQ, a self-report measure of stress states. In this experiment, the short version of the DSSQ was used (see Helton, 2004; Matthews et al., 2005). To illustrate, the Engagement scale consisted of items such as "My attention was directed towards the task", the Distress scale consisted of items such as "I felt tense", and the Worry scale consisted of items such as "I felt concerned about the impression I am making". The resulting Engagement, Distress, and Worry scale scores range from 0 (min) to 32 (max; 8 items scored from 0 = Definitely false to 4 = Definitely true). The standardized change scores for the three scales were calculated as:  $(\text{post-score} - \text{pre-score}) / (\text{standard deviation of the pre-score})$  (Helton, Warm, Matthews, Corcoran, & Dember, 2002).

- TLX, a self-report measure to assess workload (Hart & Staveland, 1988). The TLX is the most widely used measure of self-reported workload (see De Winter, 2014, for a review). Scores ranged from *very low* (0%) to *very high* (100%), except for the Performance item which ranged from *perfect* (0%) to *failure* (100%).
- Correct responses (%). The percentage of correct responses on the 2-back task (applies only to the MTD and HTD conditions).
- Heart rate (bpm).
- Heart Rate Variability. A time-domain (SDNN) and a frequency-domain (LF/HF ratio) measure were used. Both the SDNN and the LF/HF ratio were calculated from the NN intervals after a default artefact filter, using software by Vollmer (2015).
- Eye movements. Gaze spread (standard deviation of the gaze coordinates), dwell time (time focused on a particular area of interest [AOI]), and PERCLOS (percentage eye closure) were used to assess participants' attention levels to the road, environment and driving task.
- Concepts written down by the participants were categorized into four stages of automation (1) Information Acquisition, (2) Information Analysis, (3) Decision Selection, and (4) Action Implementation (Parasuraman, Sheridan, & Wickens, 2000), with a fifth category ('Other') for non-applicable concepts. The number of concepts and links between concepts were compared between the three runs.

The categorization of concepts into the four stages was performed by the first author. He obtained input from two Human Factors experts not involved in the present study, both of whom independently rated 173 selected concepts (a subsample, 50% the size of the main sample) from the experiment. These independent ratings were discussed and used by the first author to refine his categorization. Examples of categorized concepts are as follows: (1) "Condition sensor to look at road conditions" as

Information Acquisition, (2) “Calculate best route – traffic – distance – delays etc.” as Information Analysis, (3) “Artificial Intelligence” as Decision Selection, (4) “Mechanical Output, i.e. braking, acceleration” as Action Implementation, and (5) “MOT tax and insurance” as Other.

The links between the concepts’ stages were then counted, to create a 5 x 5 “To and From”-matrix. Links between concepts are an indicator of participants’ understanding of the cause-effect relationships (Revell & Stanton, 2016). Additionally, the number of links and the number of concepts served as indicators of the complexity of the mental model (Johnson-Laird, 2001).

- Verbal protocol analysis. Uttered statements of the participants within the 2-min intervals were transcribed, and per condition (i.e., ten 2-min intervals) visualised by means of a semantic network created with Leximancer (Smith, 2003). The three semantic networks were analysed and compared to assess participants’ situation awareness (see e.g., Grech, Horberry, & Smith, 2002; Salmon et al., 2014; for similar approaches).

In the present research the following settings were applied: First, only word-like concepts, such as ‘cars’ or ‘looking’, were identified (i.e., no name-like concepts, such as ‘BMW’ or ‘John’, were identified). Second, the ‘context block’ (i.e., a series of sentences that are assumed to have contextual coherence) was set to ‘break at paragraph’, with each paragraph containing the uttered statements during a 2-min interval. Third, word variants were merged. Fourth, because our analysis is concerned with colloquially spoken text, the ‘prose test threshold’ setting was set to 0. Fifth, “ehm”-concepts were disregarded manually from the thesaurus.

Within the Insight Dashboard (a quantitative analysis feature within Leximancer), the three different conditions (i.e., LTD, MTD, HTD) were compared regarding the concepts’



*strength* (i.e., the probability that a text belongs to a certain condition, given that this concept is present in the text, meaning the probabilities for the three conditions add up to 100%) and *relative frequency* (i.e., how frequently the concept occurs in the text for that condition).

The resulting outputs were three topical networks (one per condition) as well as a single quadrant report showing the strength and relative frequency of the 30 most prominent concepts per condition. A topical network is a two-dimensional projection of the co-occurrence between concepts, created using a linear clustering algorithm.

### **Statistical Analyses**

Comparisons between the three conditions were performed with paired  $t$  tests. A Bonferroni correction was used to account for multiple comparisons. Thus, a result was considered significant when the  $p$  value was smaller than  $.05/3$ .

## **Results**

### **Self-Report Questionnaires: DSSQ and TLX**

The results of the DSSQ showed a substantial loss of engagement with respect to the pre-task score in all three conditions (i.e., scores below zero, see Figure 4.1). Furthermore, the HTD condition yielded significantly higher distress than the LTD condition. The TLX showed significant differences between the three conditions, with relatively strong effects for Mental Demand, Performance, Effort, and Overall Workload (Figure 4.2).

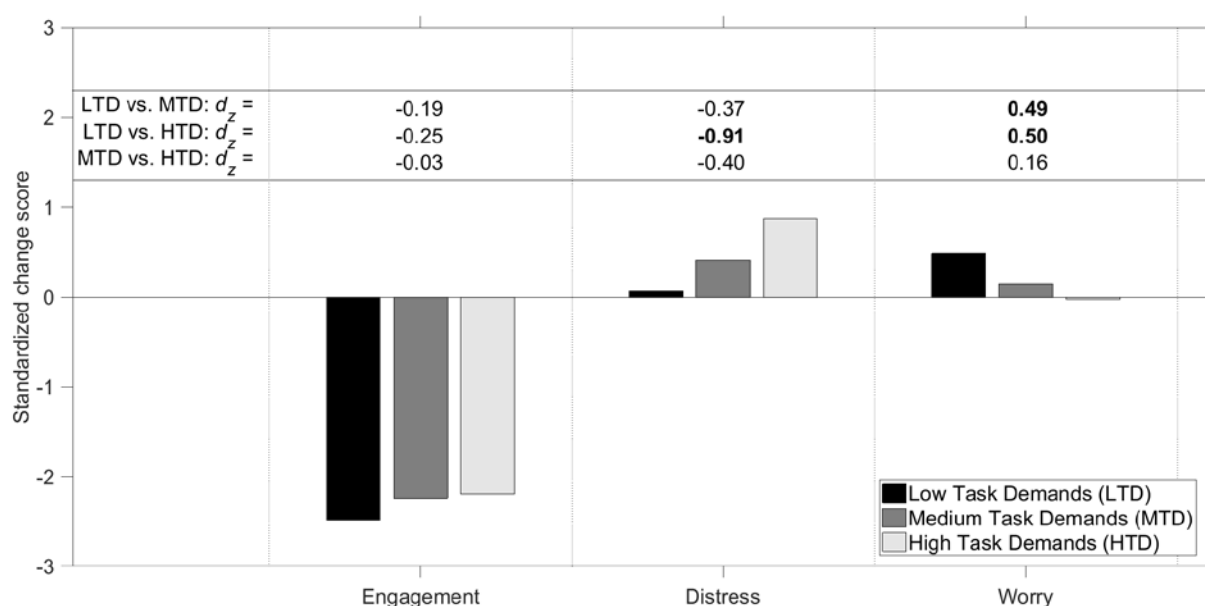


Figure 4.1. Standardized change scores for the DSSQ for the three experimental conditions. LTD = Low Task Demands, MTD = Medium Task Demands, HTD = High Task Demands. For the pairwise comparisons, the Cohen's  $d_z$  effect size is shown in boldface if  $p < .05/3$ .

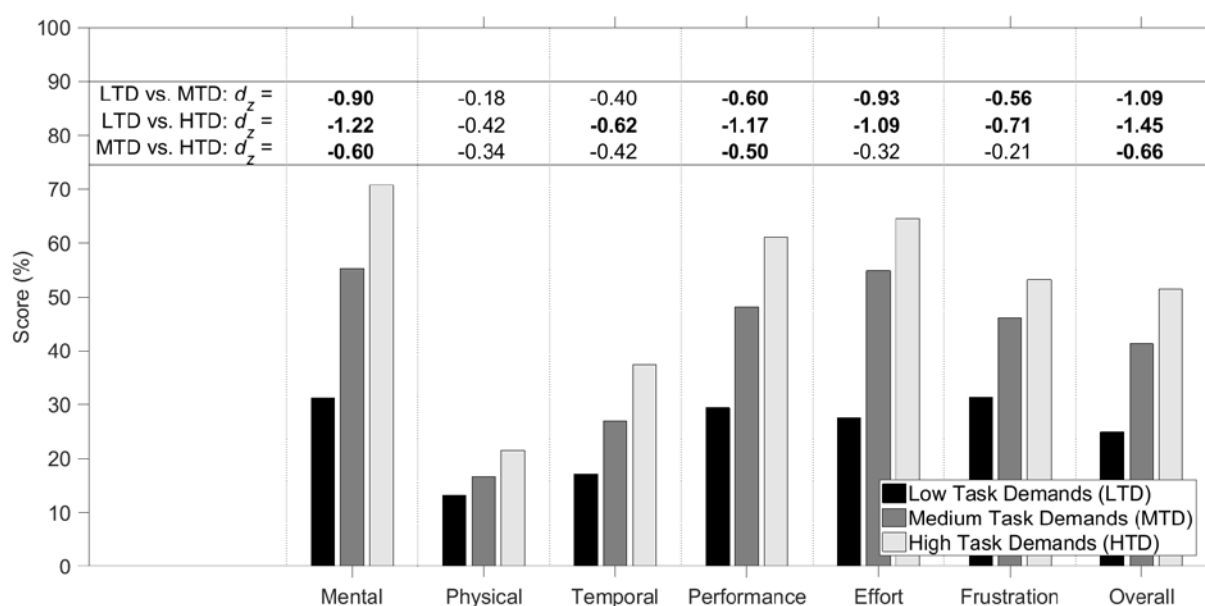


Figure 4.2. Scores on the NASA Task Load Index for the three experimental conditions. The scores are expressed as a percentage and range from Very low (0%) to Very high (100%) for the Mental Demand, Physical Demand, Temporal Demand, Effort, and Frustration items, and from Perfect (0%) to Failure (100%) for the Performance item. For the pairwise comparisons, the Cohen's  $d_z$  effect size is shown in boldface if  $p < .05/3$ .

### Performance on the 2-Back Task

The mean ( $SD$ ) percentage correctly reported letters for the MTD and HTD conditions was 41.4% (22.2%) and 64.3% (20.9%), respectively. Pairwise comparison revealed a

significant difference between the two conditions:  $t(32) = -5.66$ ,  $p < .001$ . Furthermore, a decline in task performance over time occurred (Figure 4.3).

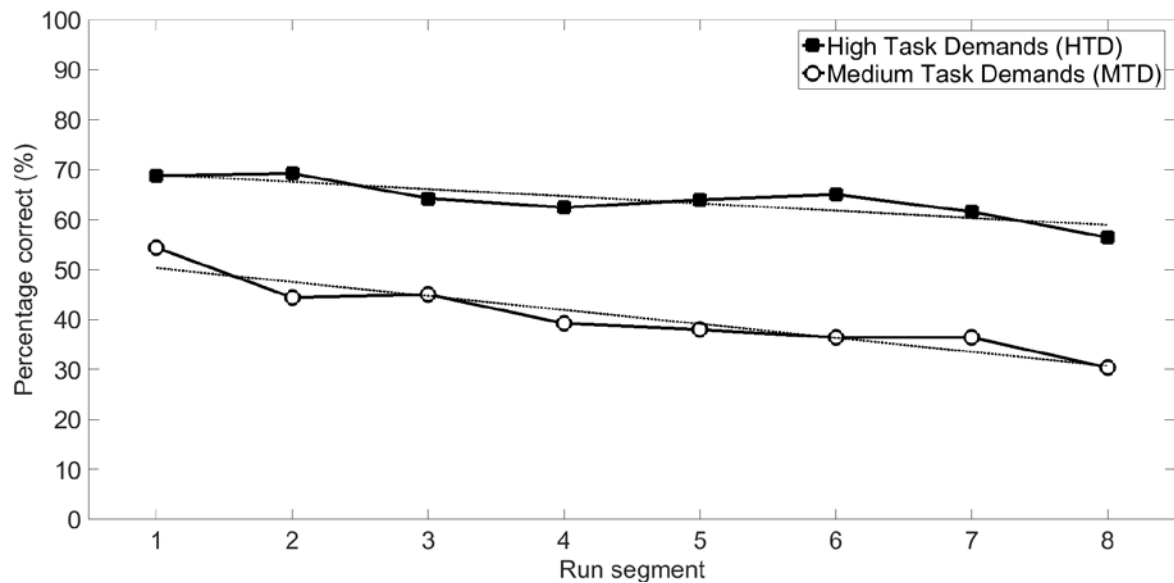


Figure 4.3. Percentage of letters reported correctly in of the 2-back task per 5-minute segment during the run. The dotted lines represent linear trend lines.

### Heart Rate

Due to data recording errors, cardiovascular data were unavailable for 31 of 99 runs (i.e., 33 participants \* 3 runs). The analysis of heart rate and heart rate variability were performed for the available 68 runs.

The mean (*SD*) heart rate for the LTD, MTD, and HTD conditions were 74.6 bpm (10.3), 75.5 bpm (9.6), and 76.4 bpm (10.9), respectively. Pairwise comparisons showed that the three conditions were not significantly different: LTD vs. MTD:  $d_z = -0.13$ ,  $p = .576$ ; LTD vs. HTD:  $d_z = -0.41$ ,  $p = .098$ ; MTD vs. HTD:  $d_z = 0.02$ ,  $p = .947$ .

A subsequent analysis on run number revealed clear differences between Run 1 ( $M = 78.2$ ,  $SD = 10.1$  bpm), Run 2 ( $M = 75.2$ ,  $SD = 10.3$  bpm), and Run 3 ( $M = 72.8$  bpm,  $SD = 9.6$  bpm). Pairwise comparisons showed significant differences: Run 1 vs. Run 2:  $d_z = 1.05$ ,  $p < .001$ ; Run 1 vs. Run 3:  $d_z = 1.50$ ,  $p < .001$ ; Run 2 vs. Run 3:  $d_z = 0.76$ ,  $p = .004$ . The run effect of heart rate is shown in Figure 4.4.

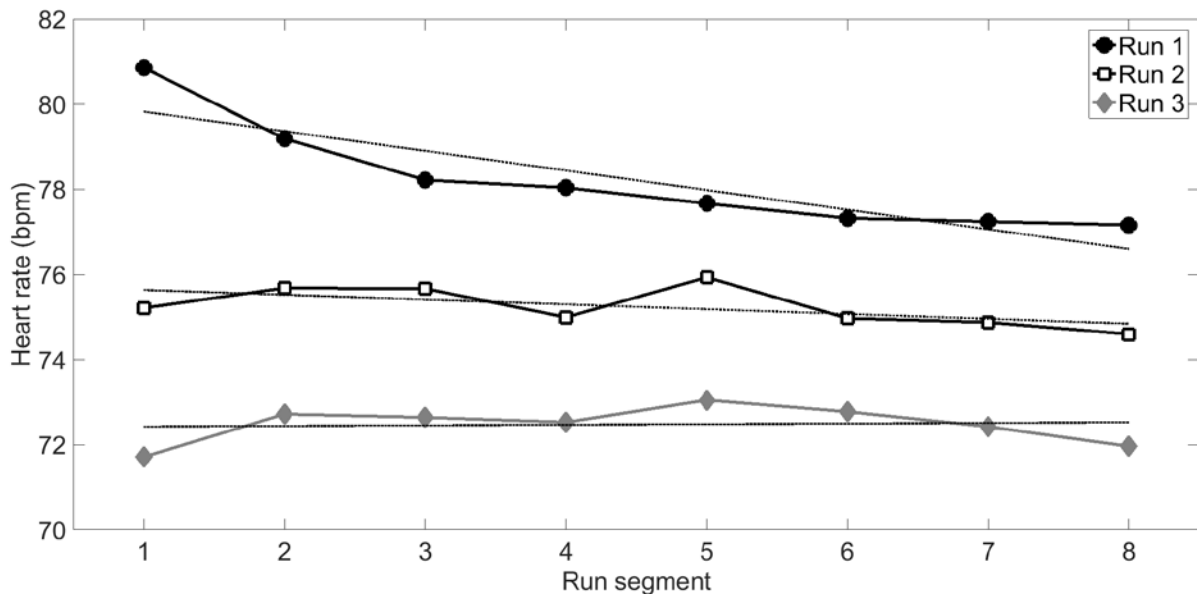


Figure 4.4. Mean heart rate during Runs 1, 2 and 3 per 5-minute segment. The dotted lines represent linear least-squares trend lines.

### Heart Rate Variability

The mean (*SD*) SDNN for the LTD, MTD, and HTD conditions was 73.5 ms (34.9), 67.0 ms (29.0), and 64.5 ms (29.4), respectively. These effects were in the expected direction, with heart rate variability being lower for higher task demands (see Figure 4.5), but pairwise comparisons showed no statistically significant differences (with Bonferroni correction) between conditions: LTD vs. MTD:  $d_z = 0.11$ ,  $p = .651$ ; LTD vs. HTD:  $d_z = 0.62$ ,  $p = .018$ ; MTD vs. HTD:  $d_z = 0.21$ ,  $p = .363$ .

The mean (*SD*) LF/HF ratio for the LTD, MTD, and HTD conditions were 1.21 (0.27), 1.35 (0.35), and 1.38 (0.40), respectively. These differences were also in the expected direction with higher task demands corresponding to a higher ratio, but were not statistically significant: LTD vs. MTD:  $d_z = -0.38$ ,  $p = .128$ ; LTD vs. HTD:  $d_z = -0.43$ ,  $p = .083$ ; MTD vs. HTD:  $d_z = -0.28$ ,  $p = .231$ .

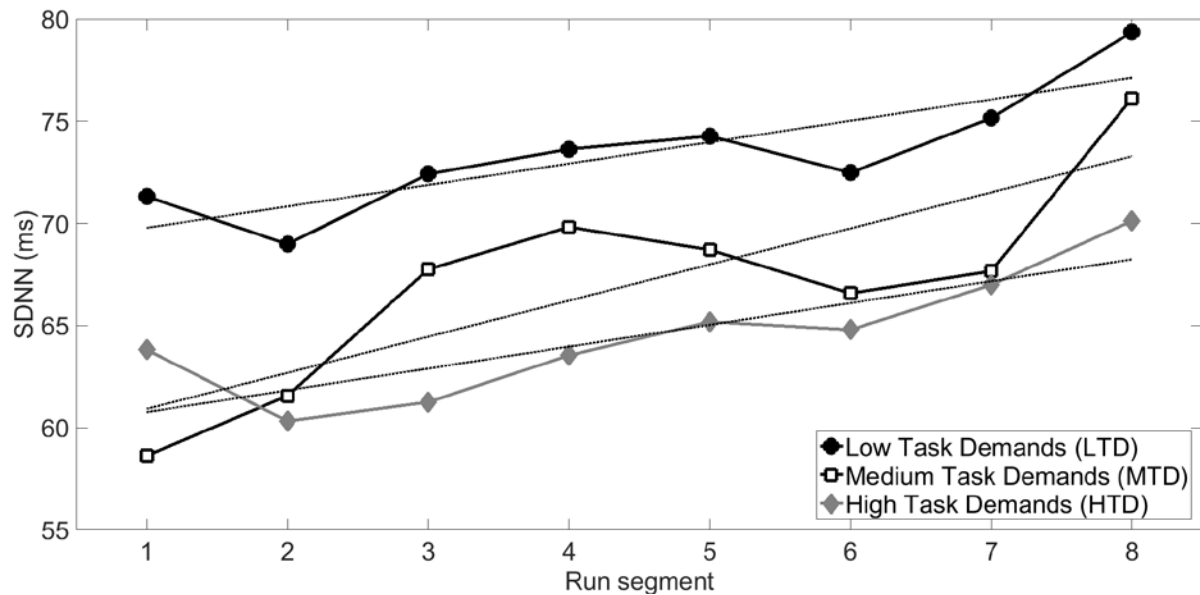


Figure 4.5. Mean SDNN for each condition per 5-minute segment. The dotted lines represent linear least-squares trend lines.

### Eye Movements

A quality check of the eye tracker data revealed that for many participants there were drifts in the eye-gaze coordinates, presumably caused by slipping of the eye tracker on the participant's head. In addition, eye movement data were often noisy or unavailable. For these reasons, we refrained from quantitative analyses of measures such as dwell time, eye closure, or fixation duration between the three task conditions.

However, in more qualitative terms, a visual inspection of the raw data revealed that participants in all three task conditions predominantly focused on the road ahead, and occasionally glanced to the mirrors or dashboard. An illustration for one run of one participant (Run 2, MTD condition) is provided in Figures 4.6 and 4.7. Figure 4.6 shows that this participant focused on the road ahead for a large portion of the time (A), and sometimes glanced into the right mirror (B), the left mirror (C), the dashboard (D), or the rear-view mirror (E).

We performed an analysis of the horizontal gaze spread (standard deviation of the horizontal gaze coordinate) for 20 participants who did not exhibit excessive noise or missing values. The results showed no significant differences between the three conditions ( $p > 0.2$  for

the three combinations). In other words, the MTD and HTD conditions did not appear to cause evident visual tunnelling as compared to the LTD condition.

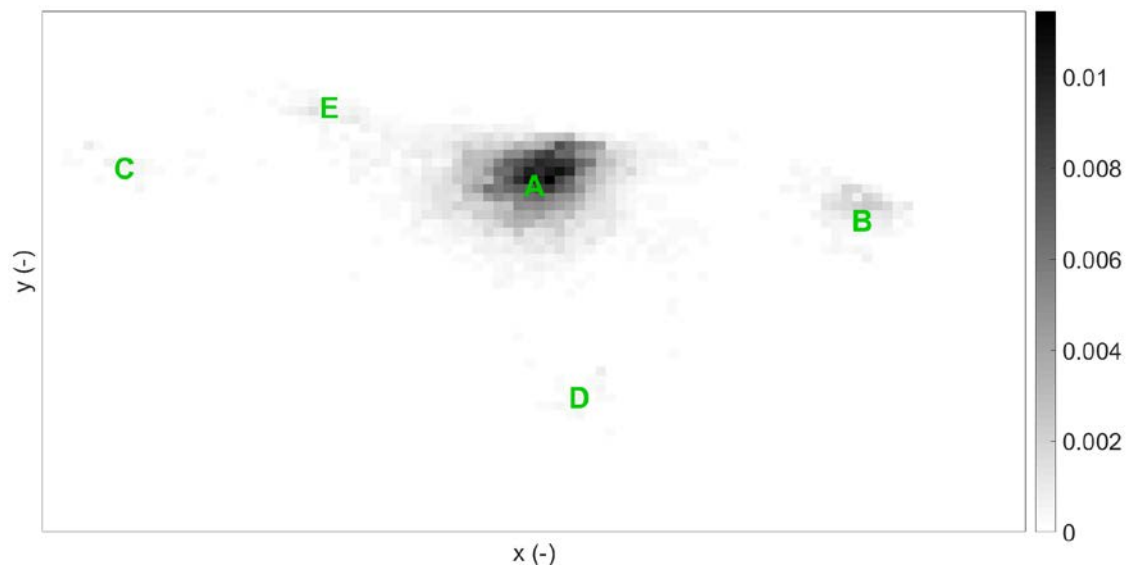


Figure 4.6. Heatmap of x- and y-coordinates of eye gaze. The darkness of the pixel indicates how frequently the participant looked at this area (darker is more frequent). The total area of all pixels in the figure equals 1. The letters illustrate the approximate locations of the road ahead (A), the right mirror (B), the left mirror (C), the dashboard (D), and the rear-view mirror (E).

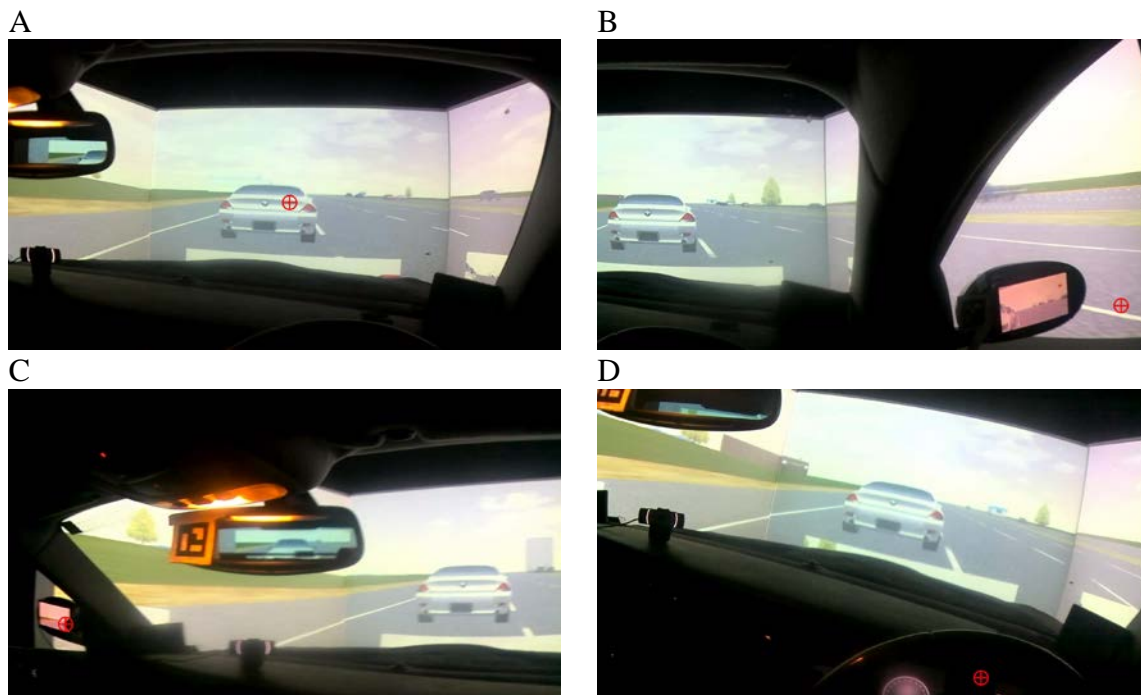


Figure 4.7. Illustrative screenshots of the head-mounted forward-facing camera of the eye tracker. The red crosshair indicates the participant's momentary gaze. A = participant glances to the car ahead, B = participant glances right of the right mirror, C = participant glances into the left mirror, D = participant glances to the dashboard. The letters A, B, C, and D correspond to the letters in Figure 4.6.

### **Mental Models based on the Quick Association Check (QuACk)**

An example of a mental model created with the QuACk method is provided in Figure 4.8. In this case, the participant produced 11 concepts and 22 links between concepts.

Noteworthy is that none of the participants changed their mental model completely at any point during the experiment. From a possible 99 (33 participants \* 3 runs) times, participants changed (added or removed links/concepts, or altered the layout) their mental model 69 times.

The mean (*SD*) number of *concepts* in the participants' baseline mental model was 8.24 (3.39), and increased to 9.42 (3.39), 10.67 (3.35), 11.15 (3.62) after Runs 1, 2, and 3, respectively. Pairwise comparisons showed significant differences between all combinations: Baseline vs. Run 1:  $d_z = -0.92, p < .001$ ; Baseline vs. Run 2:  $d_z = -1.48, p < .001$ ; Baseline vs. Run 3:  $d_z = -1.40, p < .001$ ; Run 1 vs. Run 2:  $d_z = -1.20, p < .001$ ; Run 1 vs. Run 3:  $d_z = -1.15, p < .001$ ; Run 2 vs. Run 3:  $d_z = -0.52, p = .006$ .

The mean (*SD*) number of *links* in the participants' baseline mental model was 11.45 (9.33), and rose to 13.42 (9.31), 16.21 (10.12), and 17.36 (10.49), in Runs 1, 2, and 3, respectively. Pairwise comparisons revealed significant differences between all combinations: Baseline vs. Run 1:  $d_z = -0.90, p < .001$ ; Baseline vs. Run 2:  $d_z = -1.13, p < .001$ ; Baseline vs. Run 3:  $d_z = -1.18, p < .001$ ; Run 1 vs. Run 2:  $d_z = -0.76, p < .001$ ; Run 1 vs. Run 3:  $d_z = -0.84, p < .001$ ; Run 2 vs. Run 3:  $d_z = -0.51, p = .007$ . It was further observed that the mean number of links and the mean number of concepts per participant was strongly correlated (Spearman's  $\rho = .65, N = 33$ ).

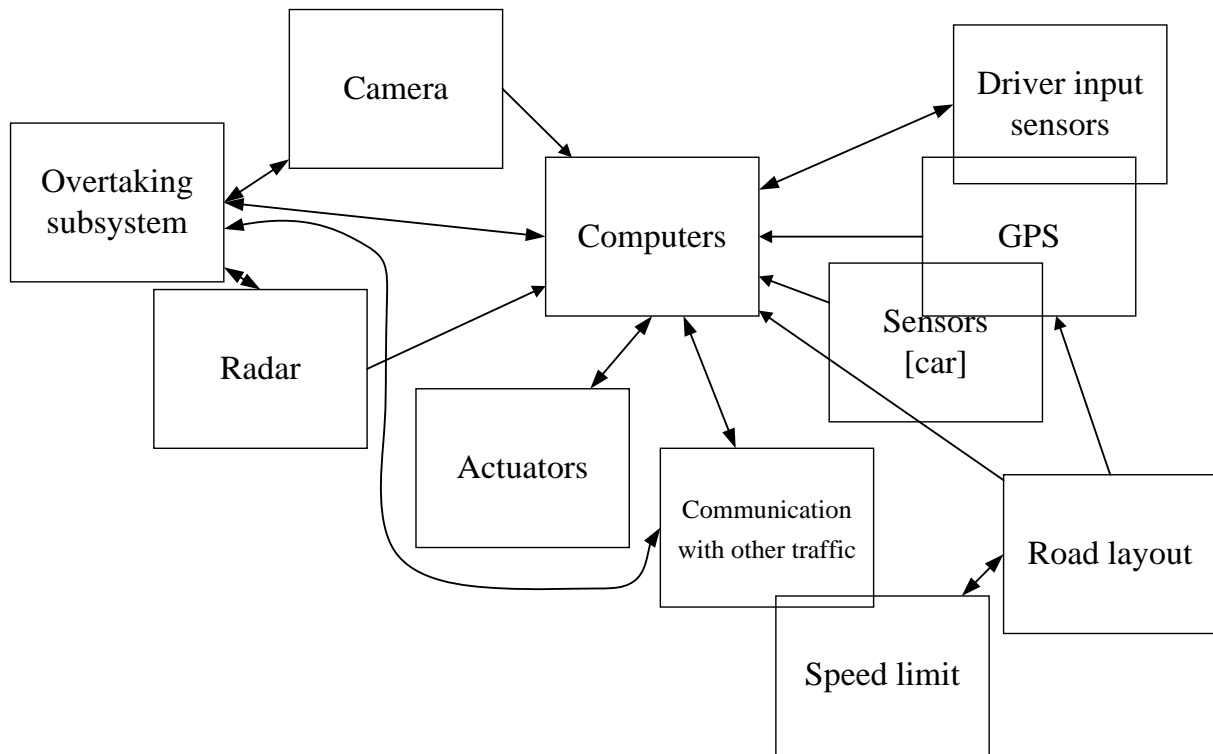
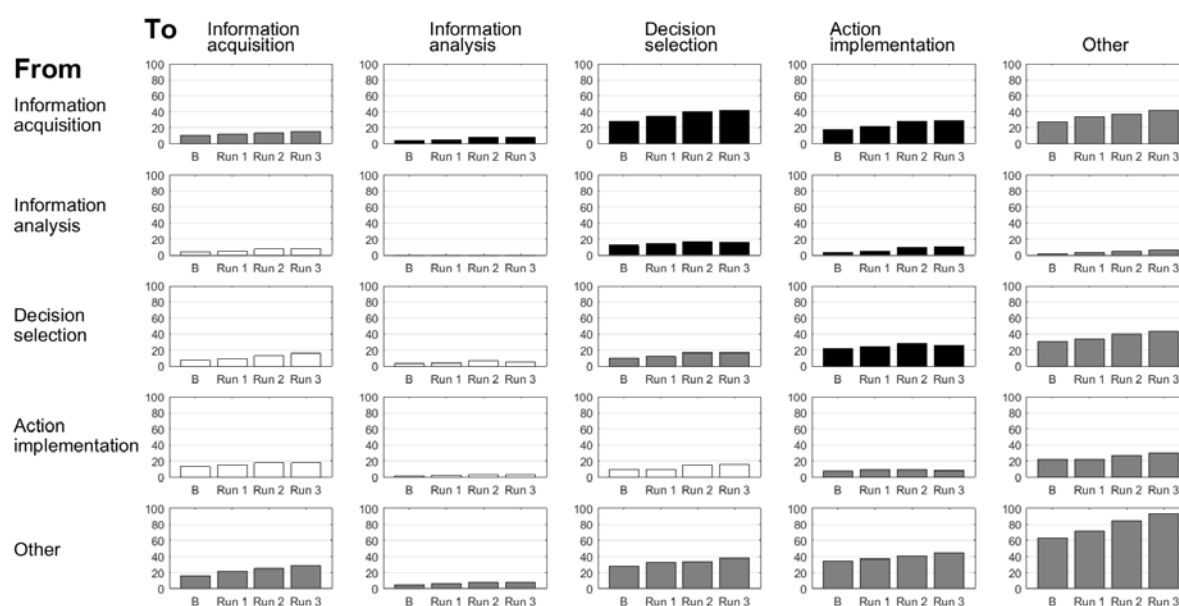


Figure 4.8. An example of a participant's mental model after Run 3 (redrawn based on how the participant actually positioned the Post-It notes and arrows on the A3 sheet).

Figure 4.9 shows the results of the participants' mental models as categorized according to the stages of automation. The majority of the links between concepts involve Other concepts, whereas the least common links involve Information Analysis concepts. Furthermore, the drawn links were more often in agreement with the order of 'stages of automation' postulated by Parasuraman et al. (2000, black bars in Figure 4.9) than in disagreement with that order (white bars in Figure 4.9).





*Figure 4.9.* Results of participants' mental models categorised into the four stages of automation (as defined by Parasuraman et al., 2000) for Baseline (B), Run 1, Run 2, and Run 3. Each bar represents the mean number of links per participant from a stage (rows) to another stage (columns). Bar graphs in black correspond to the order of the four stages (i.e., 1. Information acquisition → 2. Information analysis → 3. Decision selection → 4. Action implementation) as defined by Parasuraman et al. (2000). Bar graphs in white correspond to links that follow the opposite direction as the four-stage model by Parasuraman et al. Bar graphs in gray correspond to and from the Other concepts, and links to the same stage of automation.

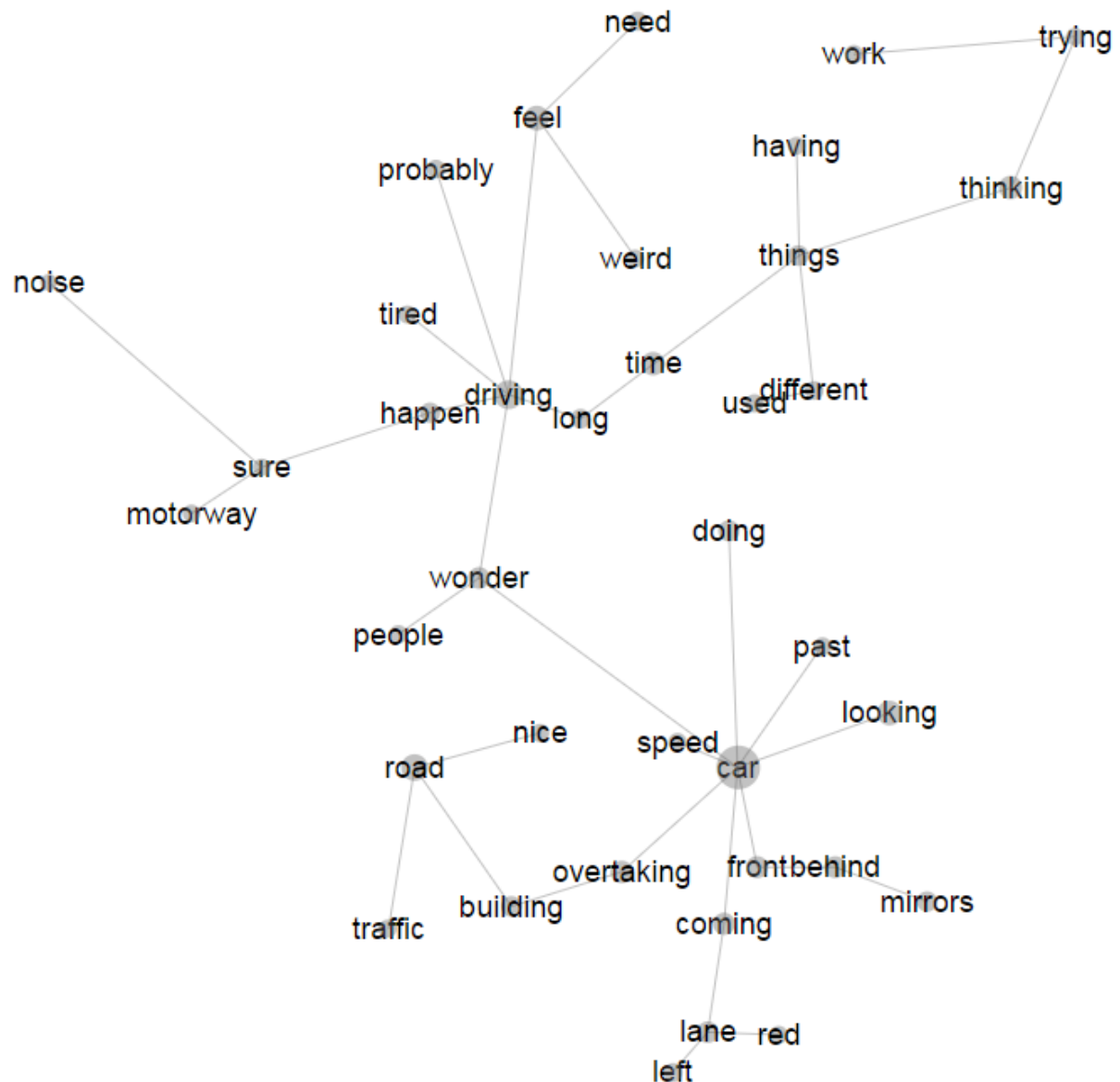
### Verbal Protocol Analysis

Of a total of 990 (33 participants x 3 runs x 10 intervals) possible 2-min intervals, 50 intervals were unavailable due to recording errors. Of the available intervals, 895 intervals contained relevant information (i.e., no untranscribable utterances, completely silent intervals, or merely containing a single word). The total number of ranked concepts for 895 intervals combined during the LTD, MTD, and HTD conditions was 2755, 2194, and 1640, respectively. The letters uttered by the participants as part of the 2-back task were not taken into account.

Figure 4.10 shows the topical networks of the LTD, MTD, and HTD conditions, respectively. It can be seen that the concepts within the statements uttered by the participants were predominantly about the car in the LTD condition and predominantly about the letters (of the 2-back task) in the HTD condition.

These observations are supported by the quadrant report (Figure 4.11), from which it is evident that in the MTD condition (blue), and particularly in the HTD condition (red), participants were occupied with trying to remember letters. The LTD condition (green) shows a relatively strong (towards the top) and frequent (towards the right) occurrence for situation and driving-task related concepts such as 'driving', 'road', 'car', 'overtaking', 'motorway', 'behind', 'front', 'lane', 'looking', whereas the strength and frequency of these concepts is comparatively low in the HTD condition. For example, in the LTD condition, the 'overtaking' concept had a strength of 48% and relative frequency of 3%. The corresponding strength and relative frequency for the MTD condition were 30% and 2%, respectively. For the HTD condition, the strength and relative frequency were 21% and 2%.

Figure 4.10a. Low Task Demand



Figures 4.10a-c. Topical networks developed through Leximancer for the LTD, MTD and HTD conditions, respectively. The greater a concept node's diameter, the higher the relative frequency of the concept within the text. The links (grey lines) indicate concepts that are strongly connected.

*Figure 4.10b. Medium Task Demand.*

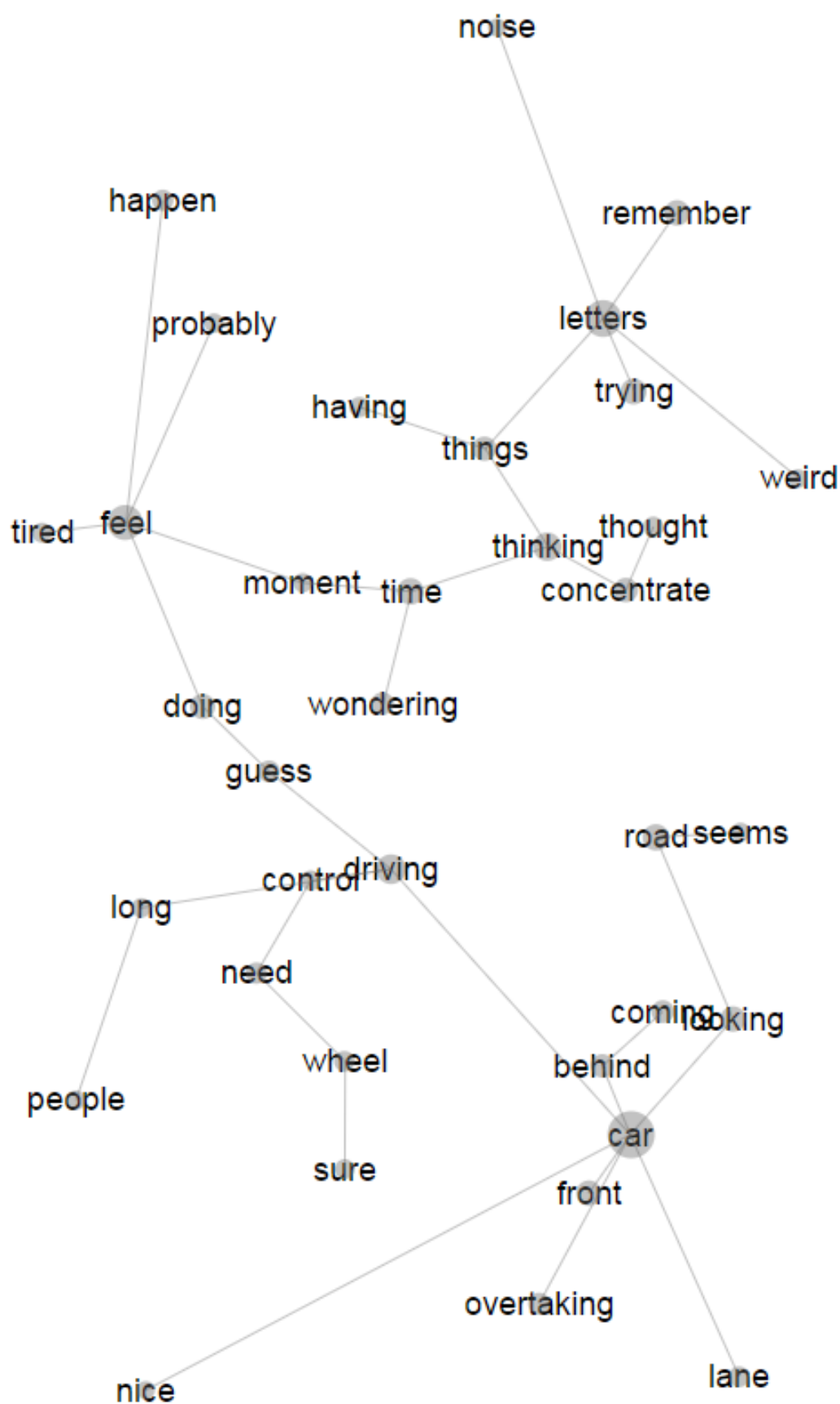
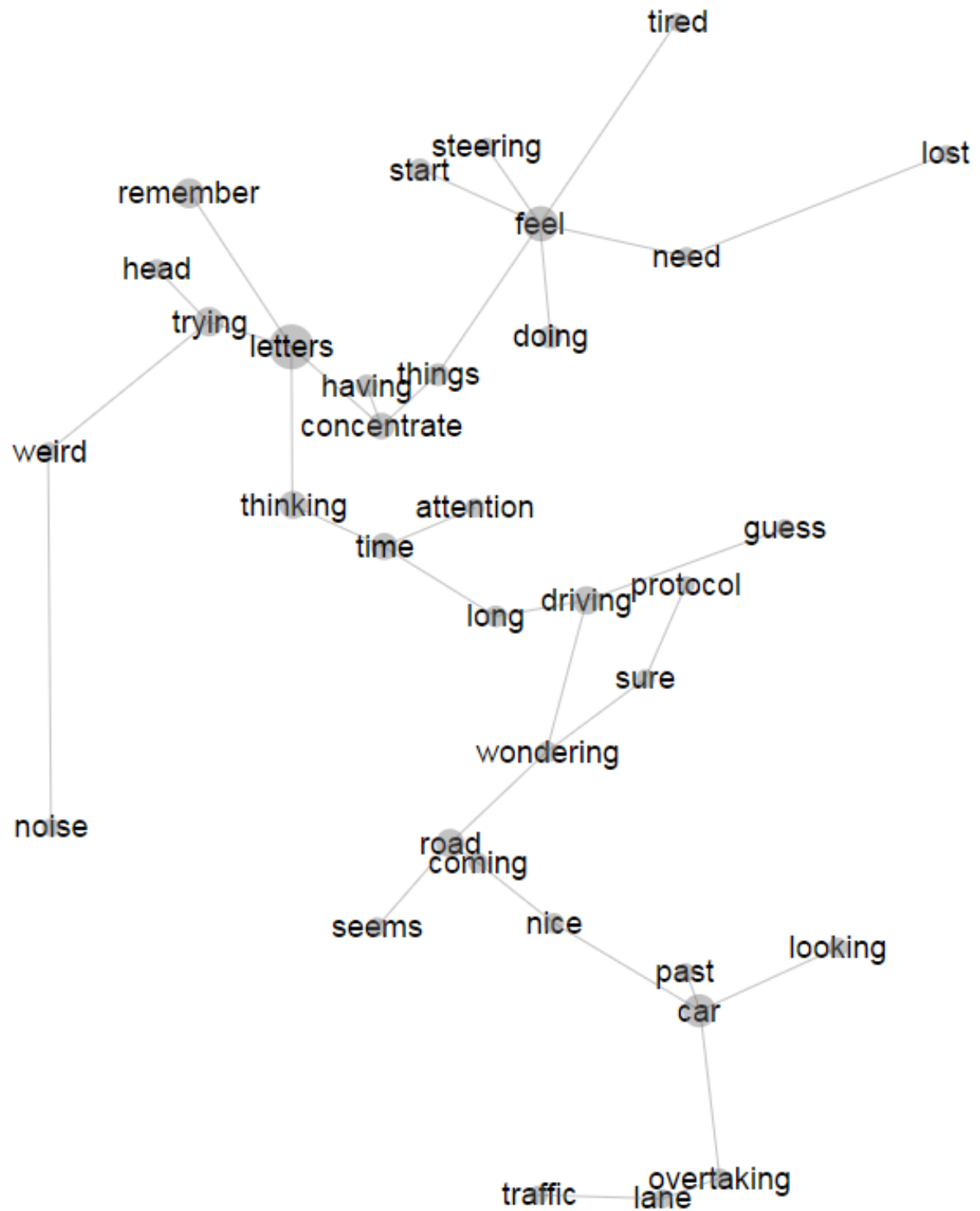


Figure 4.10c. High Task Demand.



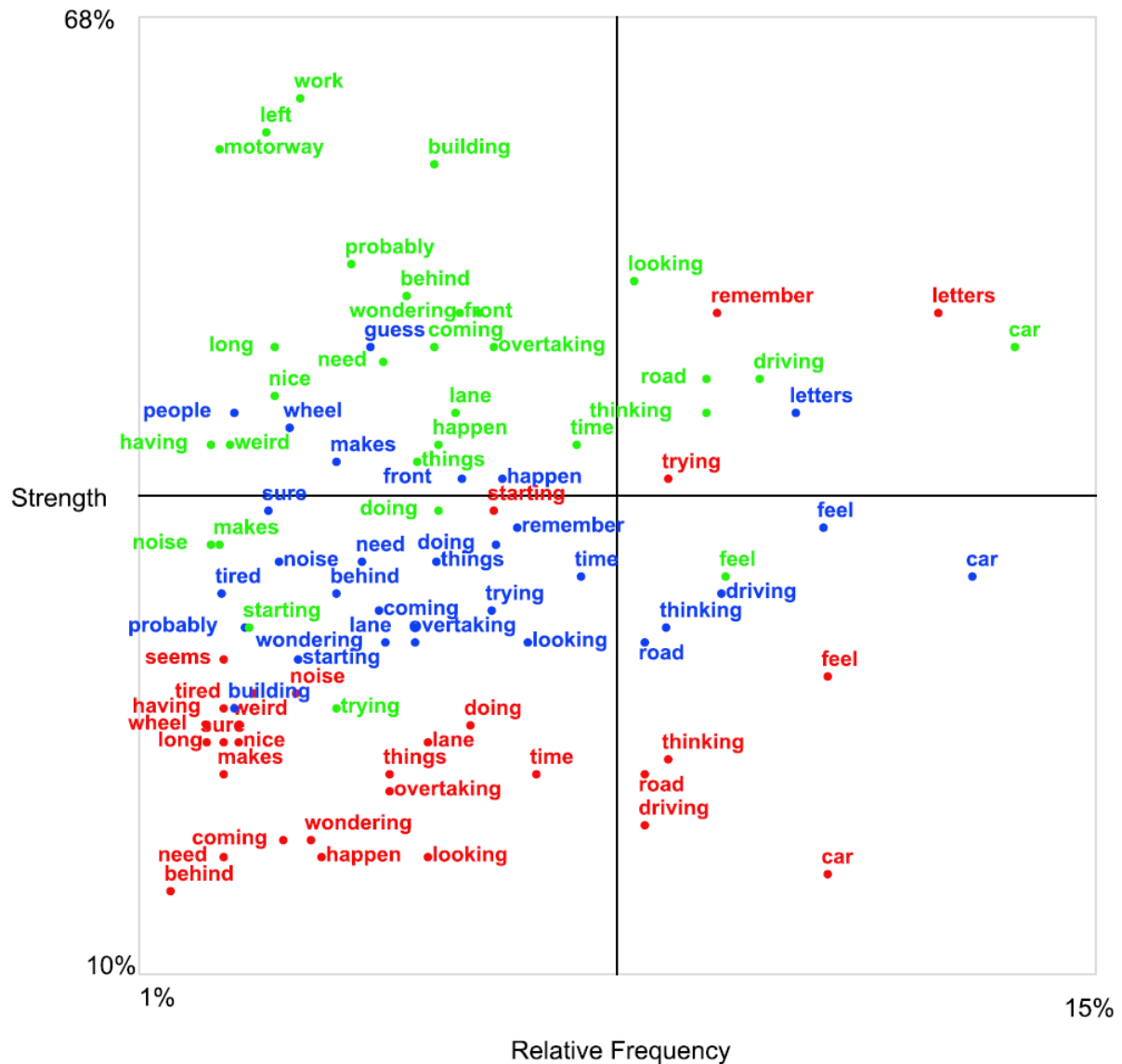


Figure 4.11. Quadrant report of the verbal protocol analysis performed with Leximancer for each of the three conditions. Green = LTD; Blue = MTD; Red = HTD. The top 30 occurring concepts per condition are displayed and placed according to its relative frequency (x-axis) and strength (y-axis) in percentage.

## Discussion

### Assessing the Effects of Mental Demands

The present study aimed to assess driver's situation awareness as a function of mental demands during automated platooning. Additionally, the development of drivers' mental models of automated driving was investigated. Participants were transported in a simulated platoon and were requested to monitor the road and intervene whenever a critical situation occurred. In two of the three conditions, participants were either required (HTD) or requested

(MTD) to perform a 2-back task by means of reporting the letter that was displayed two letters before by a pre-recorded voice through a speaker.

On a scale from *Very low* to *Very high*, participants in the present experiment reported Mental Demands of 55% (MTD) and 71% (HTD), compared to 26% (voluntary task) and 39% (visual detection task) in a previous platooning experiment by Heikoop et al. (2017). The fact that the 2-back task yielded a percentage of about 65% of correct answers is another indication that the 2-back task was indeed mentally demanding. Additionally, although the eye tracking data were not of high quality, it was clear that participants were attentive to the road in all three conditions (i.e., they were not predominantly engaging in visual secondary tasks). Furthermore, participants had a mean heart rate of 76.4 bpm in the HTD condition (a typical resting rate), and reported low levels of engagement as compared to the pre-task measurements.

Collectively, these findings illustrate that our experimental design was successful in eliciting mental demands: the mental task was not subjectively engaging, physiologically stressful, or visually distracting, yet was able to create three distinct levels of mental workload (as shown by the TLX).

### **Situation Awareness**

The verbal protocol analysis showed a clear effect of mental task demands on situation awareness. Statements regarding the remembering of letters were strong and frequent during the MTD and HTD conditions, and statements regarding the driving situation were strong and frequent in the LTD condition. Moreover, from Figure 4.11 it can be seen that participants reported to be looking around in the LTD condition, whereas this was less evident in the MTD and HTD conditions, which could be indicative of the ‘look-but-failed-to-see’ phenomenon (Hills, 1980). Another finding is that with increased mental task demands the participants uttered fewer statements. A logical explanation is that the participants had to utter responses

to the 2-back task; these 2-back responses were not taken into account in the verbal protocol analysis. A second explanation is that the participants were mentally occupied by the N-back task, thereby not having enough resources left to establish their situation awareness and utter corresponding statements about their thoughts.

Our findings add to the literature in that mental demands impair situation awareness (Ma & Kaber, 2005; Rogers et al., 2011). Although drivers may be able to counter fatigue by performing a verbal task (Atchley et al., 2014), this does not imply they remain aware of the situation around them. In fact, our results indicate that having no additional task demands is best for maintaining situation awareness, as driving related statements such as looking and driving were most prevalent in the LTD condition. Accordingly, policy makers and designers of technology should be aware that the mere recommendation for drivers to engage in a verbal task, or not engage in visually demanding tasks (e.g., working, interacting with a smartphone) is insufficient to keep drivers situationally aware; mental demands alone also reduce situation awareness.

It should be noted that the observed effects were particularly strong for the ‘strength’ dimension of the verbal protocol analysis; effects were less clear for the ‘relative frequency’ dimension. This can be explained by the fact that participants in the LTD condition uttered substantially more words than in the MTD and HTD conditions. It is likely that the secondary task (that is, to report letters) interfered directly with the verbal protocol.

Also, one could wonder whether situation awareness on the level of looking and driving (i.e., level 1 situation awareness; Endsley, 1995) would be sufficient for taking over manual control. In a non-critical take-over situation, such as during exiting the platoon, this might suffice. However, in a critical or more complex situation a higher level of situation awareness is important to act appropriately (cf. Radlmayr, Gold, Lorenz, Farid, & Bengler, 2014). Thus, it appears that multiple levels of situation awareness are necessary to be



maintained by a driver of an automated vehicle. Therefore, it is suggested that at least some level of mental demand (e.g., the level of mental demand requested during the MTD condition) should be requested from the driver, in order for him/her to maintain a higher level of situation awareness. During the MTD condition, participants uttered statements regarding guessing and remembering next to statements regarding looking and driving, an indication of a balance between the three levels of situation awareness.

### **Mental Models**

A relatively novel approach was used for the assessment of mental models, namely by means of the Quick Association Check (QuACk; Revell & Stanton, 2016). Two noteworthy findings were obtained:

First, participants did appear to have a rudimentary understanding of how automation works because links between concepts were more often in agreement with the order of the *four stages of automation* (as defined by Parasuraman et al., 2000) than in disagreement with it. Even so, participants produced highly different mental models. For example, some participants did not draw sensors or computers, but focused only on the vehicle's basic components, such as the engine, tyres, and gears (classified as 'Other' in Figure 4.9).

Second, during the course of the experiment, none of the participants overhauled their mental model completely. Participants appeared to stick to their original mental model and gradually added concepts and links, resulting in an increasingly complex mental model with run number. This may be explained by the fact that participants did not receive (dis)confirmatory information during the experimental runs: The automation always worked flawlessly, and hence participants may have had no incentive to alter their mental models, allowing for time to think about related concepts to be added. Our findings are different from Beggiano and Krems (2013), who found that non-experienced problems with the automation tend to disappear from participants' mental models. The difference between our approach and

that of Beggiato and Krems, however, is that in our case participants were not informed about problems that may occur with the automation; the mental models were entirely self-constructed.

Previous driving simulator research by Kazi, Stanton, Walker, and Young (2007) concurs that drivers tend to stick to their formed mental models of an automated driving system. In their experiment, participants were provided with a manual on the workings of an adaptive cruise control (ACC) system, as well as a list of features/functions of the ACC, and were subjected to either reliable, unreliable, or semi-reliable ACC over a ten-day period. The authors concluded that “conceptual models were consolidated over a short period of time, however they did not match that of designers’ model of Adaptive Cruise Control, thus better design solutions may be warranted.”

In summary, our results showed that mental models (operationalized via self-created concept maps) are *not* self-correcting, but rather become increasingly complex with time. This suggests that without prior information or training on automated driving systems, drivers could retain an inaccurate mental model (see also Kazi et al., 2007).

A limitation of our method is that the categorisation of concepts into four stages was often ambiguous. An example is “Computer”, which was classified as Decision Selection by the lead researcher, but which can also be plausibly classified as Information Analysis. Accordingly, the reproducibility of the results in Figure 4.9 deserves further investigation. Another limitation is that the present study was conducted among a university population. It is likely that mental models of the general population, who may be less technology-oriented than the present university sample, may be less in agreement with Parasuraman et al.’s (2000) four stages of automation.

### **Time-on-Task Effects**

This experiment showed that participants' heart rate dropped during the course of the experiment. Furthermore, a declining trend in the percentage correct answers on the 2-back task occurred. These findings suggest that participants may have become fatigued and gradually lost their vigilance. Overall, the heart rate differed more between Run numbers than between the three task demands conditions.

### **Measurement Issues**

Although the heart rate variability measures showed effects in the expected direction (i.e., lower SDNN and higher LF/HF ratio with increasing mental demands), the effects were neither strong nor statistically significant. These observations indicate that physiological indexes are not as discriminative between mental workload conditions as self-reports. One of the issues is that heart rate itself strongly correlated with SDNN (Spearman  $\rho = -0.46$  in the present experiment,  $N = 29$ ) as well as with the LF/HF ratio ( $\rho = 0.37$ ), which raises questions about the independency of these cardiovascular measures. Mehler, Reimer, and Wang (2011) previously found that heart rate itself was better able in detecting differences in both low and high workload scenarios than measures of heart rate variability.

The eye tracking data revealed several problems regarding quantitative analysis due to movement and slipping of the eye tracker during the experiment. Nevertheless, a qualitative analysis showed no significant differences between the three conditions on the account of visual tunnelling. We suggest that future research should encompass an ergonomic design of the eye tracker to avoid excessive slippage and movement of the eye tracker in order to improve data quality. Alternatively, a high-quality remote eye tracker rather than a head-mounted eye tracker may be considered.

### **Further Research**

The present study was concerned with drivers' psychological state; participants did not actually have to implement a response. Ultimately, safety is determined by behaviour, not by psychological state. Accordingly, we recommend that future research examines how drivers respond in safety-critical situations. If drivers behave unsafely when leaving a platoon, appropriate human-machine interfaces and training/instruction procedures may need to be developed to counteract this problem. Some previous research has already investigated driver behaviour after leaving a platoon. For example, studies showed increased driving speeds and decreased time headway during manual driving after having driven in a platoon (Brandenburg & Skottke, 2014; Levitan et al., 1998; Skottke, Debus, Wang, & Huestegge, 2014). Additionally, it has been found that manual drivers' headway and self-reported stress is affected when they drive next to a platoon (Gouy, Wiedemann, Stevens, Brunett, & Reed, 2014; Larburu, Sanchez, & Rodriguez, 2010).

The verbal protocol approach taken in this study could be further developed (e.g., by using a non-verbal mental task) in order to determine the different levels of situation awareness more precisely. Also concurrent psychophysiological measures could be used that are known to measure (levels of) situation awareness. For example, electroencephalography (EEG) could be used to relate fatigue patterns to situation awareness (e.g., French, Clarke, Pomeroy, Seymour, & Clark, 2007). Future research could also investigate which level of situation awareness is required or appropriate for different driving tasks or events. For example, future research could investigate what level of situation awareness is needed while transferring into and out of a platoon (i.e., intervening in a critical situation, or exiting the platoon, in mild or heavy traffic).

## Conclusion

This experiment showed that mental demands of the 2-back task have a strong effect on driver's self-reported mental demands but not on their psychophysiological responses. Driver situation awareness (as analysed by a topical network through Leximancer) was impaired due to the additional mental demands.

Furthermore, clear time-on-task effects were seen in psychophysiological measures, secondary task performance (2-back performance), and the complexity of self-constructed mental models. This suggests that not only the type of task, but also time-on-task should have a role in future research on Human Factors in automated driving. Future research should concern on-road platooning experiments in which drivers have to resume manual control.

## References

- Atchley, P., Chan, M., & Gregersen, S. (2014). A strategically timed verbal task improves performance and neurophysiological alertness during fatiguing drives. *Human Factors: The Journal of Human Factors and Ergonomics Society*, 56, 453-462. doi: 10.1177/0018720813500305
- Axelsson, J. (2017). Safety in vehicle platooning: A systematic literature review. *IEEE Transactions on Intelligent Transportation Systems*, 18, 1033-1045. doi: 10.1109/TITS.2016.2598873
- Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour*, 18, 47-57. doi: 10.1016/j.trf.2012.12.006
- Bergenheim, C., Pettersson, H., Coelingh, E., Englund, C., Shladover, S., & Tsugawa, S. (2012). Overview of platooning systems. *Proceedings of the 19th Intelligent Transportation Systems World Congress*.
- Biester, L. (2008). *Cooperative automation in automobiles*. (Doctoral dissertation), Humboldt-University, Berlin, Germany.
- Brandenburg, S., & Skottke, E-M. (2014). Switching from manual to automated driving and reverse: Are drivers behaving more risky after highly automated driving? *Proceedings of the IEEE International Conference on Intelligent Transportation Systems (ITSC)* (pp. 2978-2983). Qindao, China.
- Bueno, M., Dogan, E., Hadj Selem, F., Monacelli, E., Boverie, S., & Guillaume, A. (2016). How different mental workload levels affect the take-over control after automated driving. *IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. doi: 10.1109/ITSC.2016.7795886
- Carsten, O., Lai, F., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control task substitution in semi-automated driving: Does it matter what aspects are automated? *Special*

- Section: Human Factors and Automation in Vehicles, 54, 747-761. doi: 10.1177/0018720812460246*
- Cha, D. (2003). Driver workload comparison among road sections of automated highway systems. SAE Technical paper (No. 2003-01-0119). doi: 10.4271/2003-01-0119
- De Waard, D., Van der Hulst, M., Hoedemaeker, M., & Brookhuis, K. A. (1999). Driver behavior in an emergency situation in the Automated Highway System. *Transportation Human Factors, 1*, 67-82. doi: 10.1207/sthf0101\_7
- De Winter, J. C. F. (2014). Controversy in human factors constructs and the explosive use of the NASA TLX: A measurement perspective. *Cognition, Technology & Work, 16*, 289-297. doi: 10.1007/s10111-014-0275-1
- Endsley, M. R. (1988). Situation Awareness Global Assessment Technique (SAGAT). *Proceedings of the National Aerospace and Electronics Conference* (pp. 789-795): New York: IEEE.
- Endsley, M. R. (1995). Towards a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32-64. doi: 10.1518/001872095779049543
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-251. doi: 10.1037/0033-295X.87.3.215
- Gold, G., Berisha, I., & Bengler, K. (2015). Utilization of drivetime - Performing non-driving related tasks while driving highly automated. *Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting, 59*, 1666-1670. doi: 10.1177/1541931215591360
- Gouy, M., Wiedemann, K., Stevens, A., Brunett, G., & Reed, N. (2014). Driving next to automated vehicle platoons: How do short time headways influence non-platoon drivers' longitudinal control? *Transportation Research Part F: Traffic Psychology and Behaviour, 27*, 264-273. doi: 10.1016/j.trf.2014.03.003
- Grech, M. R., Horberry, T., & Smith, A. (2002). Human error in maritime operations: Analyses of accident reports using the Leximancer tool. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting, 46*, 1718-1721. doi: 10.1177/154193120204601906
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology, 52*, 139-183. doi: 10.1016/S0166-4115(08)62386-9
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2016). Psychological constructs in driving automation: A consensus model and critical comment on construct proliferation. *Theoretical Issues in Ergonomics Science, 17*, 284-303. doi: 10.1080/1463922X.2015.1101507
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2017). Effects of platooning on signal-detection performance: A driving simulator study. *Applied Ergonomics, 60*, 116-127. doi: 10.1016/j.apergo.2016.10.016
- Helton, W. S. (2004). Validation of a short stress state questionnaire. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, 48*, 1238-1242. doi: 10.1177/154193120404801107
- Helton, W. S., Warm, J. S., Matthews, G. S., Corcoran, K. J., & Dember, W. N. (2002). Further tests of an abbreviated vigilance task: Effect of signal salience and jet aircraft noise on performance and stress. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 46*, 1546-1550. doi: 10.1177/154193120204601704
- Hills, B. L. (1980). Vision, visibility, and perception in driving. *Perception, 9*, 183-216. doi: 10.1068/p090183
- Johnson-Laird, P. N. (2001). Mental models and deduction. *TRENDS in Cognitive Science, 5*, 434-442. doi: 10.1016/S1364-6613(00)01751-4

- Kazi, T. A., Stanton, N. A., Walker, G. H., & Young, M. S. (2007). Designer driving: Drivers' conceptual models and level of trust in adaptive cruise control. *International Journal of Vehicle Design*, 45, 339-360. doi: 10.1504/IJVD.2007.014909
- Kunze, R., Haberstroh, M. A., Hauck, E., Ramakers, R., Henning, K., & Jeschke, S. (2011). Automated truck platoons on motorways – A contribution to the safety on roads. In S. Jeschke, I. Isenhardt & K. Henning (Eds.), *Automation, Communication and Cybernetics in Science and Engineering 2009/2010*: Springer: Berlin, Heidelberg.
- Kyriakidis, M., Happee, R., & De Winter, J. C. F. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, 32, 127-140. doi: 10.1016/j.trf.2015.04.014
- Larburu, M., Sanchez, J., & Rodriguez, D. J. (2010). Safe road trains for environment: Human factors' aspects in dual mode transport systems. *Proceedings of the 17th World Congress on Intelligent Transport Systems*. Busan, Korea.
- Larson, J., Liang, K. Y., & Johansson, K. H. (2015). A distributed framework for coordinated heavy-duty vehicle platooning. *IEEE Transactions on Intelligent Transportation Systems*, 16, 419–429. doi: 10.1109/TITS.2014.2320133
- Leviton, L., Golembiewski, G., & Bloomfield, J. R. (1998). Human Factors issues for Automated Highway Systems. *ITS Journal – Intelligent Transportation Systems Journal: Technology, Planning, and Operations*, 4, 21-47. doi: 10.1080/10248079808903735
- Llaneras, R. E., Salinger, J., & Green, C. A. (2013). Human Factors issues associated with limited ability autonomous driving systems: Drivers' allocation of visual attention to the forward road. *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (pp. 92-98). Bolton Landing, NY.
- Louw, T., Madigan, R., Carsten, O., & Merat, N. (2016). Were they in the loop during automated driving? Links between visual attention and crash potential. *Injury Prevention*. doi: 10.1136/injuryprev-2016-042155
- Ma, R., & Kaber, D. B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics*, 35, 939-953. doi: 10.1016/j.ergon.2005.04.002
- Matthews, G. S., Emo, A. K., & Funke, G. J. (2005). *A short version of the Dundee Stress State Questionnaire*. Paper presented at the Twelfth Meeting of the International Society for the Study of Individual Differences, Adelaide, Australia.
- Mehlenbacher, B., Wogalter, M. S., & Laughery, K. R. (2002). On the reading of product owners' manuals: Perceptions and product complexity. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, 730-734. doi: 10.1177/154193120204600610
- Mehler, B., Reimer, B., & Wang, Y. (2011). A comparison of heart rate and heart rate variability indices in distinguishing single-task driving and driving under secondary cognitive workload. *Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 590–597). Lake Tahoe, CA.
- Neubauer, C., Matthews, G. S., & Saxby, D. (2014). Fatigue in the automated vehicle: Do games and conversation distract or energize the driver? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58, 2053-2057. doi: 10.1177/1541931214581432
- Nilsson, J. (2014). *Safe transitions to manual driving from faulty automated driving system*. (Doctoral dissertation), Chalmers University of Technology, Gothenburg, Sweden.

- Omae, M., Hashimoto, N., Sugamoto, T., & Shimizu, H. (2005). Measurement of driver's reaction time to failure of steering controller during automatic driving. *Review of Automotive Engineering*, 26, 213-215.
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, 30, 286-297. doi: 10.1109/3468.844354
- Petermeijer, S. M., Cieler, S., & De Winter, J. C. F. (2017). Comparing spatially static and dynamic vibrotactile take-over requests in the driver seat. *Accident Analysis and Prevention*, 99, 218-227. doi: 10.1016/j.aap.2016.12.001
- Ploeg, J., Van de Wouw, N., & Nijmeijer, H. (2014). Lp string stability of cascaded systems: Application to vehicle platooning. *Proceedings of the IEEE Transactions on Control Systems Technology*, 22, 786-793. doi: 10.1109/TCST.2013.2258346
- Ren, W., & Green, D. (1994). Continuous platooning: A new evolutionary and operating concept for automated highway systems. *Proceedings of the American Control Conference*, 11(20), 21-25. doi: 10.1109/ACC.1994.751685
- Revell, K. M. A., & Stanton, N. A. (2012). Models of models: Filtering and bias rings in depiction of knowledge structures and their implications for design. *Ergonomics*, 55, 1073-1092. doi: 10.1080/00140139.2012.692818
- Revell, K. M. A., & Stanton, N. A. (2016). The Quick Association Check (QuACk): A resource-light, 'bias robust' method for exploring the relationship between mental models and behaviour patterns with home heating systems. *Theoretical Issues in Ergonomics Science*, 17, 554-587. doi: 10.1080/1463922X.2016.1180439
- Rogers, M., Zhang, Y., Kaber, D., & Liang, Y. (2011). The effects of visual and cognitive distraction on driver situation awareness. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics. EPCE 2011. Lecture Notes in Computer Science* (pp. 6781). Berlin, Heidelberg: Springer.
- Salmon, P. M., Lenne, M. G., Walker, G. H., Stanton, N. A., & Filtness, A. (2014). Exploring schema-driven differences in situation awareness between road users: An on-road study of driver, cyclist and motorcyclist situation awareness. *Ergonomics*, 57, 191-209. doi: 10.1080/00140139.2013.867077
- Sarter, N. B., & Woods, D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology*, 1, 45-57. doi: 10.1207/s15327108ijap0101\_4
- Saxby, D. J., Matthews, G. S., & Neubauer, C. (2017). The relationship between cell phone use and management of driver fatigue: It's complicated. *Journal of Safety Research*, 61, 129-140. doi: 10.1016/j.jsr.2017.02.016
- Shaffer, F., & Combatalade, D. C. (2013). Don't add or miss a beat: A guide to cleaner Heart Rate Variability recordings. *Biofeedback*, 41, 121-130. doi: 10.5298/1081-5937-41.3.04
- Skottke, E-M., Debus, G., Wang, L., & Huestegge, L. (2014). Carryover effects of highly automated convoy driving on subsequent manual driving performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56, 1272-1283. doi: 10.1177/0018720814524594
- Smith, A. E. (2003). Automatic extraction of semantic networks from text using Leximancer. *Proceedings of the HLT-NAACL 2003 Demonstrations* (pp. 23-24). Edmonton.
- Verwey, W. B., & Zaidel, D. M. (1999). Preventing drowsiness accidents by an alertness maintenance device. *Accident Analysis and Prevention*, 31, 199-211. doi: 10.1016/S0001-4575(98)00062-1



- Vollmer, M. (2015). A robust, simple and reliable measure of heart rate variability using relative RR intervals. *IEEE Conference on Computing in Cardiology Conference (CinC)*, 609-612. doi: 10.1109/CIC.2015.7410984
- Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors: The Journal of Human Factors and Ergonomics Society*, 44, 365-375. doi: 10.1518/0018720024497709
- Young, M. S., & Stanton, N. A. (2007). What's skill got to do with it? Vehicle automation and driver mental workload. *Ergonomics*, 50, 1324-1339. doi: 10.1080/00140130701318855

*Appendix A.* Extensive Ranked Concept Lists for each condition. Only the concepts that make up the Topical Network of Figure 4.11 are included in the list. The TOTAL count is the word count when analysing all three conditions together. Note: when using Leximancer, this results in slightly different word counts per condition than when simply adding the three separate lists, as the calculation of a concept inclusion is amongst others based on the average amount of sentences per block, which is different for each condition, hence also for the three combined. Relevance is the percentage of occurrence of a concept relative to the most occurring concept. Therefore, the most occurring concept is 100%, regardless of its occurrence count.

Ranked Concept List LW			Ranked Concept List MW			Ranked Concept List HW		
Word-Like	Count	Relevance	Word-Like	Count	Relevance	Word-Like	Count	Relevance
car	352	100%	car	242	100%	letters	165	100%
driving	165	47%	feel	161	67%	feel	121	73%
feel	150	43%	letters	147	61%	car	121	73%
thinking	144	41%	driving	119	49%	remember	87	53%
road	141	40%	thinking	102	42%	trying	78	47%
looking	119	34%	road	95	39%	thinking	76	46%
time	97	28%	time	76	31%	road	70	42%
overtaking	77	22%	looking	70	29%	driving	69	42%
wonder	74	21%	remember	67	28%	time	49	30%
front	70	20%	doing	63	26%	start	46	28%
lane	69	20%	trying	61	25%	concentrate	43	26%
doing	66	19%	front	56	23%	doing	41	25%
building	65	18%	things	52	21%	lane	37	22%
things	62	18%	happen	52	21%	looking	35	21%
coming	60	17%	lane	51	21%	things	34	21%
behind	60	17%	concentrate	49	20%	overtaking	32	19%
happen	59	17%	overtaking	49	20%	wondering	26	16%
mirrors	59	17%	wondering	45	19%	noise	26	16%
need	56	16%	guess	43	18%	coming	24	15%
probably	51	14%	coming	41	17%	weird	24	15%
trying	44	12%	need	41	17%	tired	23	14%
work	43	12%	behind	40	17%	nice	22	13%
nice	41	12%	wheel	34	14%	head	22	13%
left	40	11%	noise	33	14%	having	21	13%
long	39	11%	thought	32	13%	long	21	13%
used	36	10%	sure	32	13%	seems	21	13%
weird	36	10%	probably	30	12%	traffic	21	13%
motorway	35	10%	people	29	12%	attention	20	12%
having	34	10%	tired	28	12%	sure	20	12%
noise	34	10%	nice	28	12%	need	20	12%
tired	33	09%	weird	27	11%	lost	20	12%
different	32	09%	control	26	11%	steering	18	11%
people	32	09%	having	25	10%	guess	17	10%
speed	31	09%	moment	25	10%	past	15	09%
traffic	31	09%	long	23	10%	protocol	12	07%
sure	31	09%	seems	22	09%			
past	30	09%						
red	30	09%						
total	2755	100%		2194	80%		1640	60%

## Acclimatizing to Automation

---

### Driver Workload and Stress during Partially Automated Car Following in Real Traffic

#### Abstract

Automated driving systems are increasingly prevalent on public roads, but there is currently little knowledge on the level of workload and stress of drivers operating an automated vehicle in a real environment. The present study aimed to measure driver workload and stress during partially automated driving in real traffic. We recorded heart rate, heart rate variability, respiratory rate, and subjective responses of nine test drivers in the Tesla Model S with Autopilot. The participants, who were experienced with driver assistance systems but naïve to the Tesla, completed a 32 min motorway route back and forth while following a lead car in regular traffic. In one of the two drives, participants performed a heads-up detection task of bridges they went underneath. Averaged across the two drives, the participants' mean self-reported overall workload score on the NASA Task Load Index was 19%. Moreover, the participants showed a reduction of heart rate and self-reported workload over time, suggesting that the participants became accustomed to the experiment and technology. The mean hit (i.e., pressing the button near a bridge) rate in the detection task was 88%. In conclusion, driving with the Tesla Autopilot on a motorway involved a low level of workload that decreased with time on task.

Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2017). *Acclimatizing to automation: Driver workload and stress during partially automated car following in real traffic*. Manuscript submitted for publication.

## **Introduction**

### **Workload and Stress in Automated Driving**

Cars that provide combined longitudinal and lateral automated control support have recently been introduced on the market. Automated driving may be expected to reduce workload and stress as compared to manual driving because the driver does not have to control the vehicle. However, unless the driving task is fully automated (SAE level 5), automated driving may cause high workload and stress, because the driver needs to supervise both the human-machine interface and the state of the car in relation to the outside environment (for an illustration, see Figure 5.1).

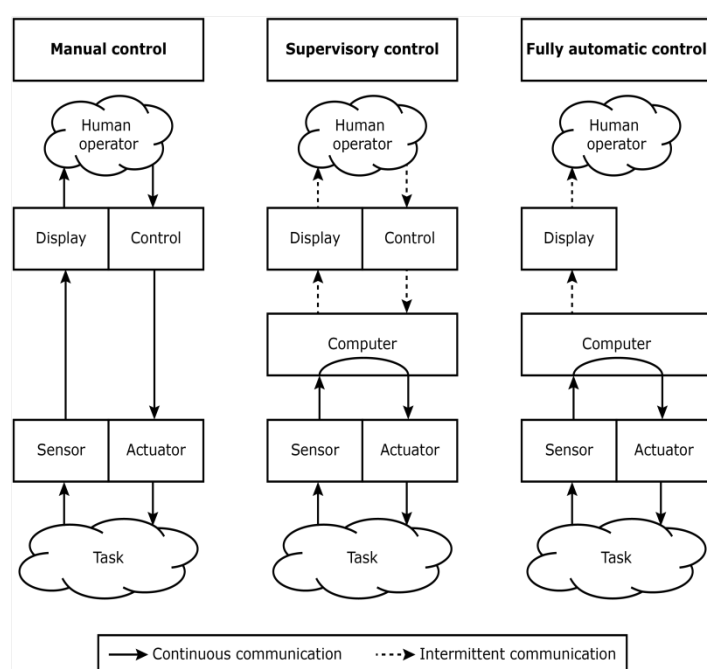
More specifically, the driver of an automated car has to remain attentive to reclaim manual control if required (Casner, Hutchins, & Norman, 2016; Stanton, Young, & McCaulder, 1997), a task that may be demanding and stressful (Hancock, 2015). Furthermore, the type of supervisory control shown in Figure 5.1 may cause out-of-the-loop problems, such as loss of situation awareness and mode errors, which resemble those observed in aviation and process control (e.g., Haslbeck & Hoermann, 2016; Kaber & Endsley, 1997; Metzger & Parasuraman, 2001; see also Stanton & Marsden, 1996). A survey by Dikmen and Burns (2016) among 121 Tesla owners found that automation failures (e.g., failure to detect lanes) were frequent but not perceived as risky. Furthermore, the majority of respondents indicated that it is important to remain alert and to be aware of the automation's limitations.

### **Prior Research on Workload and Stress in Automated Driving**

The majority of Human Factors research on driver workload in automated vehicles has been conducted in driving simulators (see De Winter, Happee, Martens, & Stanton, 2014 for a review). Overall, the results indicate that the self-reported workload as assessed with the NASA Task Load Index (TLX) is substantially lower in automated driving than in manual

driving (see De Winter et al., 2014 for a review), and on the low end of the scale from 0 to 100% (see Table 5.1 for an overview).

A small number of on-road studies are available. Recently, Endsley (2017) conducted a single-subject naturalistic driving study using her Tesla Model S over a six-month period. She reported that situation awareness increased when using automation, because less focus was needed on controlling the vehicle, and more attention could be devoted to looking at traffic and road signage. However, Endsley also experienced various issues of mode confusion and unexpected automation transitions, as well as loss of attention. Endsley further found that ratings of satisfaction, usefulness, and trust gradually increased from months 1–2 towards months 5–6, which is in line with the results of a longitudinal naturalistic driving study on adaptive cruise control (ACC) with 15 participant (Beggiato, Pereira, Petzoldt, & Krems, 2015). Additionally, overall self-reported workload was low, averaging at about 1.3 during months 1–2 and 1.0 during months 3–4, on a scale from 0 to 5 (Endsley, 2017).



*Figure 5.1.* Manual control (left), supervisory control (middle), and fully automated control (right). In manual control, the driver controls the car via manipulators (steering wheel & pedals) and continuously receives information from the car in the environment (i.e. task). In fully automatic driving, the human has no contribution to the driving task other than to set a destination (or to press an emergency stop button). Hence, the driver is taken out of the control loop completely. In supervisory control, the driver interacts with a computer that closes the control loop via sensors and actuators, while the driver intermittently (1) provides instructions to the computer, (2) receives information via displays, and (3) receives information from the car in the environment (from Sheridan, 2002).

Eriksson, Banks, and Stanton (2017) let 12 test drivers use the Tesla autopilot for about 20 minutes per participant. Participants each experienced approximately 12 automation-to-manual control transitions, and completed the NASA-TLX after the ride. The mean overall

workload was 19%. Stapel, Mullakka-Babu, and Happee (2017) conducted an on-road highway driving study in which 15 participants used the Tesla Autopilot for about 20 minutes. The authors found overall low levels of workload among participants (between 10% and 43%), with the type of road (busy city ring versus relatively empty highway) and prior experience with the Tesla Model S being moderator variables (Table 5.1). In another on-road study, Banks and Stanton (2016) tested a prototype version of automated longitudinal and lateral control in addition to a driver-initiated auto-overtaking system. These authors found relatively high workload on the NASA-TLX (median of 42%) during 9 minutes of automated driving per participant.

The discrepancy between the results of Banks and Stanton (2016) and the findings of Eriksson et al. (2017) and Stapel et al. (2017) may be caused by the fact that the prototype system tested by Banks and Stanton, which included a heads-up display and offered overtake suggestions, was difficult to use or that participants were still learning how to use it. Because the participants in Banks and Stanton (2016) drove only 9 minutes with the automation system, the high workload levels “may be a simple reflection of the fact that these ratings were collected during first time use of the automated system”, p. 393.

McDowell, Nunez, Hutchins, and Metcalfe (2008) and Davis, Animashaun, Schoenherr, and McDowell (2008) performed on-road trials with automated military convoys. In these studies, where there was no other traffic and, because they were military experiments, object detection was of primary importance. The results showed that automated driving reduced workload and improved performance in object detection in comparison to manual driving.

On-road studies may be expected to yield higher workload than simulator studies, because the latter involve no physical risk of accidents. However, in some cases, on-road studies actually yielded lower workload than simulator-based studies. For example, the

reported workload in Eriksson et al. (2017) involved experienced test drivers and did not include a secondary task; participants were merely required to take over and relinquish control of the vehicle throughout the experiment. In Manawadu et al. (2015), critical events were triggered, to which the participants had to respond.

### Aim of the Present Study

The present study aimed to assess whether on-road automated driving with the Tesla Model S alleviates driver workload over time. Both the on-road studies of Eriksson et al. (2017) and Stapel et al. (2017) consisted of approximately 20 min of highway driving with the automation engaged (excluding a familiarization drive) and did not report on temporal effects. Our study consisted of 64 minutes of automated highway driving per participant. Additionally, in our study a simple detection task was used to add extra task demands on top of the regular monitoring demands during automated driving, which is similar to the approach taken in a previous platooning experiment in a driving simulator (Heikoop, De Winter, Van Arem, and Stanton (2017). In Heikoop et al. (2017), it was found that the detection task (i.e., to detect red cars on the road) increased self-reported mental demands compared to not performing a detection task. We expected to find a similar effect in this study.

Table 5.1

*Overview of workload measurements presenting a NASA-TLX overall workload (TLX OW) score in automated driving studies.*

Reference	Simulator /road	Sample size	Mean TLX OW
Banks and Stanton (2016)	On-road	32	42% (median)
Borojeni, Chuang, Heuten, and Boll (2016)	Simulator	21	30%
Damböck, Weißgerber, Kienle, and Bengler (2013)	Simulator	24	33%
De Winter, Stanton, Price, and Mistry (2016)	Simulator	24	31% (exp. 1)
	Simulator	27	31% (exp. 2)
Eriksson et al. (2017)	On-road (Tesla)	12	19%
Eriksson and Stanton (2017)	Simulator	26	21%
Heikoop et al. (2017)	Simulator	22	28%

Large, Banks, Burnett, Baverstock, and Skrypchuk (2017)	Simulator	30	36% (partial automation), 21% (high automation)
Manawadu, Kamezaki, and Sugano (2015)	Simulator	6 (novices)	36%
	Simulator	6 (experienced)	30%
McDowell et al. (2008)	On-road (military)	11	40%
Petermeijer, Bazilinskyy, and De Winter (2017)	Simulator	24	28% (with auditory and vibrotactile feedback)
Petermeijer, Cieler, and De Winter (2017)	Simulator	18	22%, 36% (with N-Back task)
Saxby, Matthews, Warm, Hitchcock, and Neubauer (2013)	Simulator	36	34% (exp. 1)
	Simulator	56	27% (exp. 2)
Schwalk, Kalogerakis, and Maier (2015)	Simulator	24	21%
Stapel et al. (2017)	On-road (Tesla)	8 (no experience with Tesla)	25% (empty highway), 43% (city ring)
	On-road (Tesla)	7 (experienced with Tesla)	10% (empty highway), 24% (city ring)
Young (2000)	Simulator	18	23%
Young and Stanton (2004)	Simulator	12	12% (exp. 1)
	Simulator	12	12% (exp. 2)
Young and Stanton (2007)	Simulator	24 (novice drivers)	11%
	Simulator	30 (learner drivers)	13%
	Simulator	30 (expert drivers)	20%
	Simulator	30 (advanced drivers)	24%

## Methods

### Participants

Nine participants (seven males, two females) aged between 25 and 47 years ( $M = 35.44$ ;  $SD = 8.26$ ) with 6 to 30 years of self-reported driving experience ( $M = 17.56$ ;  $SD = 8.46$ ) took part in this experiment. The participants were employees of a large automotive company. Eight participants indicated that they drove every day and one participant indicated driving 4–6 days a week. Two participants indicated they drove up to 10,000 miles, five up to 20,000, one up to 30,000, and one up to 50,000 miles in the past year. All participants had completed level-2 driver training, an extended driver training specifically designed for people who drive as part of their job, and which serves as a legal requirement for insurance purposes. All participants had driven various supercars before and had experience with advanced driver



assistance systems (e.g., adaptive cruise control, lane keeping assist), but had no experience with the Tesla Autopilot. No incentive was provided to the participants, and all participants gave written informed consent. The study was approved by the Ethics Research Governance Office of the University of Southampton under submission ERGO number 19091.

### **Apparatus**

The experiment was performed with a Tesla Model S 90D with Autopilot as the participants' vehicle (PV) and a Jaguar XF as a lead vehicle (LV). The LV was used for safety reasons. With a forward-looking radar, forward-facing camera, and ultrasonic sensors, the Autopilot can steer, adjust speed, detect obstacles, and apply brakes automatically ("Full self-driving hardware on all cars," 2015). The Tesla Autopilot can be characterised as SAE J3016 level 2 automation (i.e., partial automation) because both steering and speed control are automated, and the driver is still expected to monitor the driving environment (NHTSA, 2017).

The Traffic-Aware Cruise Control (TACC) of the PV was set to 1, which was the closest following distance and which translates to a time headway of about 1 second. This headway corresponds to common headways in highway traffic (Brackstone & McDonald, 2007; Neubert, Santen, Schadschneider, & Schreckenberg, 1999; Song & Wang, 2010; Treiber, Kesting, & Helbing, 2006), and was sufficiently short to have a low likelihood of other cars merging in between the PV and LV.

Participants wore electrocardiography (ECG) equipment linked to LabChart 8 that captured their cardiovascular and respiratory activity. This ECG equipment consisted of the AD Instruments PowerLab 26T Teaching Series, three MLA2505 biopotential electrodes and lead wires with disposable ECG electrode patches, and the MLT1132 respiratory belt transducer. The electrodes were placed in a triangular configuration. For male participants, one electrode was placed over the xiphoid process, and two electrodes below the far ends of

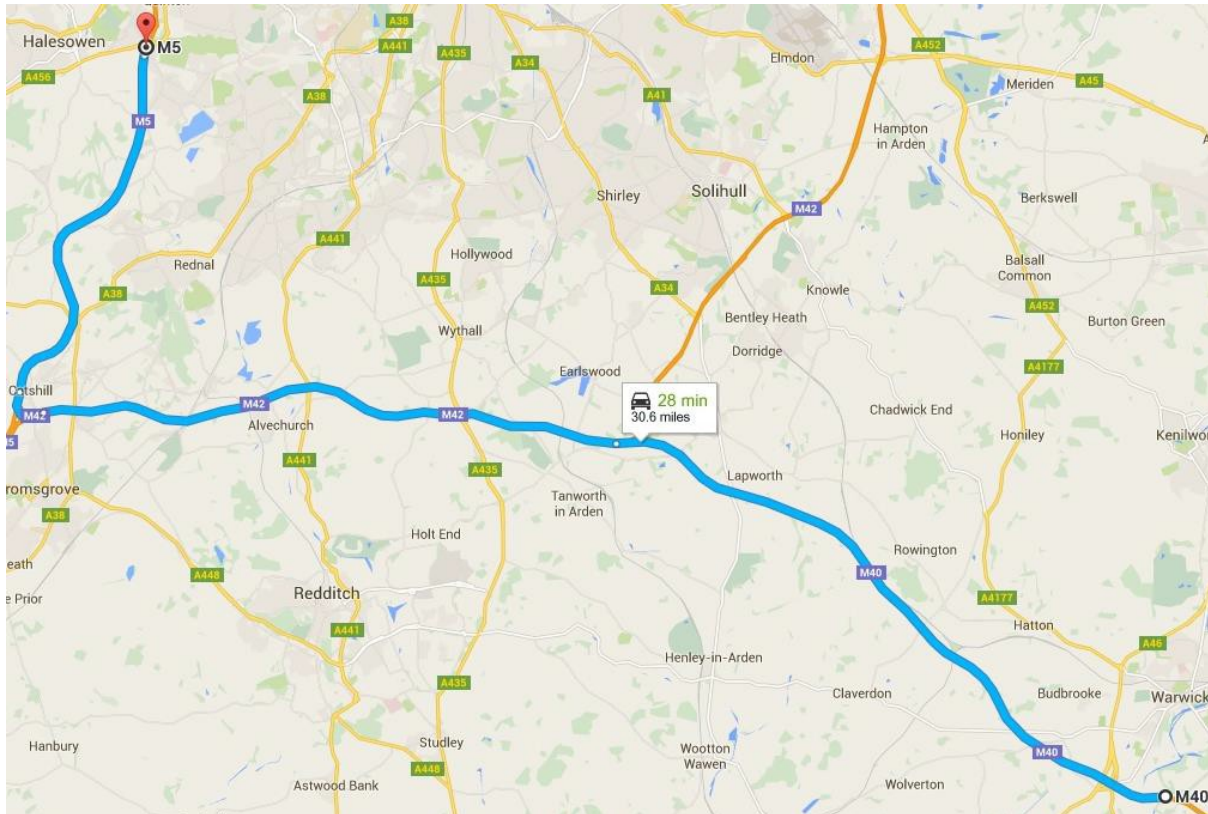
the collar bones. For female participants, one electrode was placed at the top of the sternum and two electrodes below the ribs on both sides. This gender-based distinction was mainly made for comfort purposes (see e.g., Shaffer & Combatalade, 2013). The respiratory belt was placed over the clothes around the chest.

### **Environment**

The experiment took place on March 14–18, 2016. Participants drove on the left (slow) lanes of the British dual three-lane motorways M40, M42, and M5, for which the speed limit is 70 mph (112 km/h). Participants completed two drives during daytime outside of rush hours. The first drive was completed between entry point 14 of the M40 northbound and M5 northbound exit point 3 (Figure 5.2). In the second drive, the participants drove back to the starting point. Specifically, the second drive was completed between the motorway entry point at the service stations after entry point 3 of the M5 southbound and the M40 southbound until exit point 14.

### **Procedure**

All participants received a training trial and completed two drives of approximately 32 minutes each. Vigilance research has shown that detection performance exhibits a decay function with time on task (Mackworth, 1964). Furthermore, it has been found that after 15 minutes the most substantial deterioration of detection performance has taken place (see a review by Teichner, 1974, reporting that “at least half of the final loss is completed within the first 15 min”). Because the average driving trip in Europe and the U.S. is between 20 and 30 minutes (McKenzie & Rapino, 2011; Pasaoglu et al., 2014), it may be assumed that the present study is representative of the first exposures to a new automated driving system on public roads.



*Figure 5.2.* Map (from Google Maps) displaying the northbound route, starting at entry point 14 of the M40, and ending at exit point 3 of the M5. The southbound route went in the opposite direction, starting at entry point 3 of the M5, and ending at exit point 14 of the M40.

Before the experiment, the participant performed a test drive on a test track. Upon arrival at the test track site, the participant received paper instructions explaining that he/she would be driving within a highly automated platoon. Furthermore, a consent form, a demographics questionnaire, and the pre-task Dundee Stress State Questionnaire (DSSQ) were provided. After having completed these questionnaires, the participant was taken to the passenger seat of the PV and introduced to the safety driver. The safety driver performed a lap on the test track and showcased the Autopilot, as well as several details of the car. After that lap, the participant and safety driver changed seats, and the participant drove the car until they were comfortable driving manually and with the Autopilot feature. Then the ECG electrodes were attached after which the participant drove to the selected motorway entry point, following the LV. After entering the motorway, the Autopilot was engaged by the participant, and the experiment started. The safety driver sat in the seat next to the participant, and

verbally intervened if the participant did not act appropriately or safely (e.g., when the participant did not override the automation when he/she should). The experimenter sat in the rear seat, monitoring the equipment and making notes of events during the experiment. Before the first drive, the participant was discouraged from interacting with the safety driver or the experimenter for the duration of the experiment. Thus, the interaction between the safety driver and the experimenter was kept to a minimum.

In the occasions where another vehicle merged in between the PV and the LV, the participants were instructed by the safety driver to remain in automated mode and follow this other vehicle. However, if the gap with the LV became large, then the participants were instructed to follow the LV again by overtaking the outside traffic while it was emphasised to try to remain in automated mode. An automated lane change could be performed by using the indicator stalk while holding the steering wheel. All events such as lane changes, merges, and Autopilot (dis)engagements were recorded by the experimenter using paper and pencil. Summed across the nine participants, a total of 33 and 37 lane changes (of which 16 and 21 automated) occurred for Drive 1 and 2, respectively. A manual override occurred 6 and 3 times during Drive 1 and Drive 2, respectively.

At the end of the first drive, the participants exited the motorway and stopped at a nearby parking lot. They were then provided with the post-task DSSQ and the NASA Task Load Index (TLX). Once completed, the participants followed the LV to the motorway again and performed the second drive. At the end of the second drive, the participants were again provided with the post-task DSSQ and TLX.

### **Independent Variables**

The experiment consisted of two drives, either with (DT) or without (NT) a detection task, in counterbalanced order. Specifically, five participants completed the second 'southbound' drive with the detection task, and four participants completed the first

‘northbound’ drive with the detection task. Without the detection task, participants had to follow the LV as their only objective. With the detection task, they also had to detect the bridges they went underneath by pressing a handheld button (Figure 5.3). In the first drive, participants drove underneath 50 bridges, and during the second drive, participants drove underneath 47 bridges. Photos of the bridges are available as supplementary material.

Learning/acclimatization effects were assessed by comparing the results of the first drive with the results of the second drive.



*Figure 5.3.* Photo taken during the experiment: the participant on the right presses the handheld button when detecting a bridge. The safety driver is sitting in the left (passenger) seat. The vehicle in front is the LV the participant had to follow during the experiment.

### **Dependent Measures**

The following measures were calculated per participant for each of the two drives:

- Duration of the drive (s).
- Mean speed (km/h), recorded with a GPS application on a smartphone.
- Hit rate of the bridges (% of bridges detected). The hit rate was calculated by linking the known locations of the bridges (as retrieved from Google Maps) with the locations at the moments of button presses (recorded with a GPS application on a smartphone).

An algorithm was written that matched bridges with the nearest button press in terms of radial distance, until all bridges were assigned to a button press or no button presses were left (each button press could be assigned to one bridge only). Button presses which followed each other within  $\frac{2}{3}$  seconds (i.e., accidental double pressing of the button) were discarded. Furthermore, if the nearest button press was more than 1,250 m from the bridge, then this bridge was marked as a miss. The liberal threshold of 1,250 m was used, because there were several sources of inaccuracy in the locations of the button presses. Specifically, (1) The GPS signal had a limited temporal resolution (0.2 Hz, which at an average speed of 86 km/h amounts to a travelled distance of about 120 m), (2) Some participants pressed the button late (i.e., when being beneath a bridge) while others pressed the button early (i.e., when the bridge could first be seen), and (3) The GPS recording had limited accuracy (the 50<sup>th</sup>, 95<sup>th</sup>, and 99<sup>th</sup> percentile of the estimated accuracy were 24 m, 249 m, and 965 m, respectively). By definition, the miss rate equals 100% minus the hit rate (Tanner Jr., Wilson, & Swets, 1954).

- False alarm rate (% of false alarms relative to the number of bridges). A button press was considered a false alarm when after determining the hits, there were still button presses unaccounted for. Figure 5.4 provides an illustration of the hits and false alarms for Participant 1.
- Heart rate (bpm). The heart rate was regarded as a measure of stress (Healey & Picard, 2004).
- SDNN (ms), a time-domain measure of mental workload (De Waard, 1996; Jorna, 1992; Vidulich & Tsang, 2012). The SDNN was defined as the mean of the standard deviation (SD) of all Normal to Normal peak intervals (NN) in the ECG signal per 5-min segment along the drive (SDNN index, see Task Force of the European Society of

Cardiology and the North American Society of Pacing and Electrophysiology, 1996).

A low SDNN value is interpreted as high workload (Fallahi, Motamedzade, Heidarimoghadam, Soltanian, & Miyake, 2016; see also Heikoop et al., 2017). See Figure 5.5 for an illustration of the calculation process.

- LF/HF ratio, a frequency-domain measure of mental workload. This spectral analysis of the NN interval calculates the power in the low-frequency (LF) 0.04–0.15 Hz range relative to the power in the high frequency (HF) 0.15–0.40 Hz range. A high LF/HF ratio is indicative of high workload (Cinaz, Arnrich, La Marca, & Tröster, 2013; Suriya-Prakash, John-Preetham, & Sharma, 2015). Both the SDNN and the LF/HF ratio were calculated from the NN intervals after a default NN artefact filter using an open-source MATLAB program (Vollmer, 2015).
- Respiratory rate (bpm). Because the respiratory belt transducer produced a noisy signal (presumably because of in-vehicle vibrations) and may contain drifts and other artefacts, the signal was filtered with a second-order Butterworth 0.1–1.0 Hz bandpass filter. This frequency range incorporates a typical human respiratory rate of 0.25 Hz. Next, the data were rank transformed to remove outliers, and subsequently a discrete Fourier transformation was applied to retrieve the frequency with maximum amplitude (see Figure 5.6 for illustration).
- DSSQ, a self-report measure of stress and fatigue (Matthews, Szalma, Panganiban, Neubauer, & Warm, 2013). In this experiment, version 1.3 of the DSSQ was used (Matthews, Campbell, & Falconer, 2000). Standardized change scores for each scale of the DSSQ were calculated as follows:  $(\text{post-score} - \text{pre-score}) / (\text{standard deviation of the pre-score})$  (Helton, Warm, Matthews, Corcoran, & Dember, 2002). The scores for the three scales (Engagement, Distress, and Worry) were calculated by averaging four subscales and averaging them to result in one score for each element (based on

Fairclough & Venables, 2005; Matthews, 2014; Matthews et al., 2002). Task Engagement consists of the subscales (1) Energetic Arousal, (2) Success Motivation, (3) Intrinsic Motivation, and (4) Concentration. Distress consists of (5) Tense Arousal, (6) Hedonic Tone, (7) Control & Confidence, and (8) Anger/Frustration. Finally, Worry consists of (9) Self-Focused Attention, (10) Self-Esteem, (11) Task-Relevant Interference, and (12) Task- Irrelevant Interference. We imputed missing answers (4% of the total) using the nearest-neighbour method.

- NASA-TLX, a self-report measure to assess workload (Hart & Staveland, 1988). The ‘raw’ approach was used, also known as the Raw TLX (RTLX). This approach does not apply weights to the scales (Hart, 2006).

The mean speed, duration, heart rate, SDNN, LF/HF ratio, and respiratory rate were calculated from the moment that the participant was 200 m in front of the first bridge until 200 m after the participant passed the last bridge. SDNN was calculated as the average across six available 5-min segments.



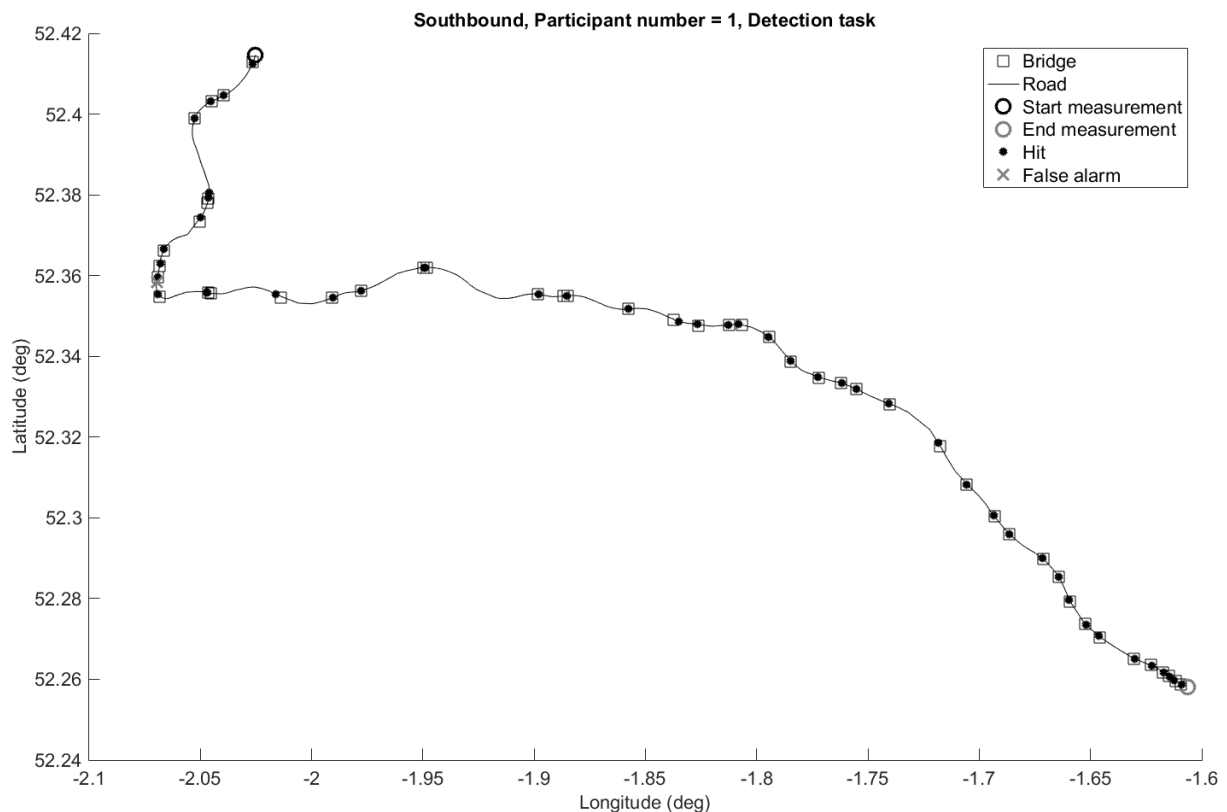
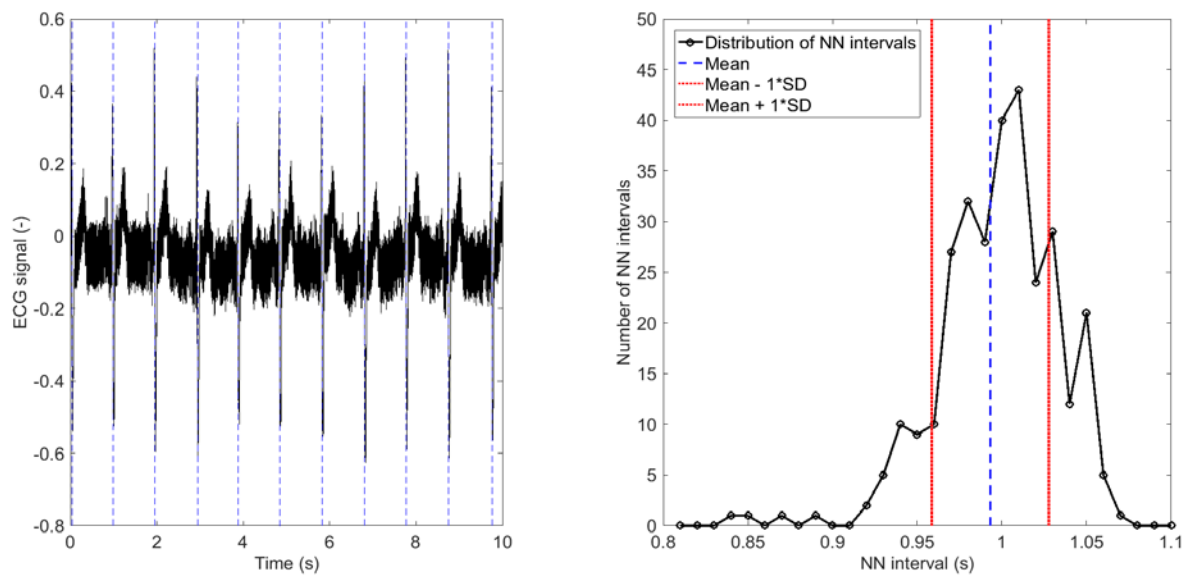


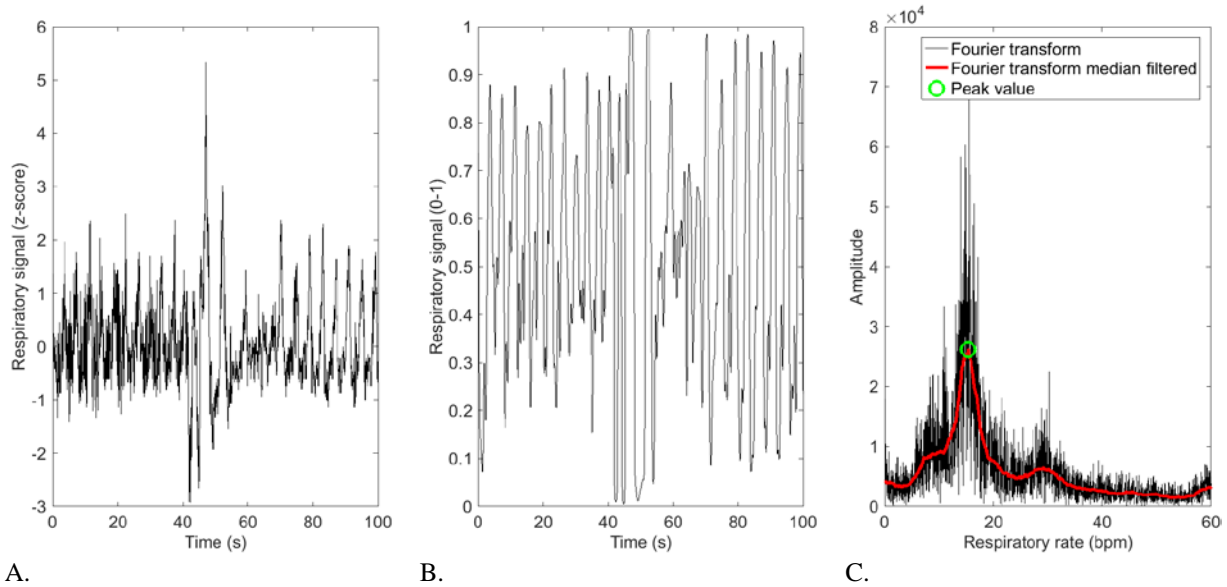
Figure 5.4. Illustration of GPS data and the detection task. In this case, the participant detected all 47 bridges (hit rate = 100%) and had 1 false alarm (false alarm rate = 2.1%). Where a single dot is visible for two bridges, two button presses appeared in the same GPS sample (the GPS recorded the position every 5 seconds).



A.

B.

Figure 5.5. Illustration of the calculation of SDNN. A) ECG signal with extracted NN intervals (first 10 s of Drive 1 of Participant #1). B) Distribution of NN intervals with the mean and standard deviation of the NN intervals (SDNN = 34.6 ms; based on first 300 s of Drive 1 of Participant #1).



A. B. C.  
*Figure 5.6.* Illustration of data processing of the respiratory signal. A) z-transformed raw signal (first 100 s of Drive 1 of Participant #1). B) Filtered signal, rank-transformed and scaled from 0 to 1 (first 100 s of Drive 1 of Participant #1). C) Discrete Fourier transform with identified peak value (based on the entire Drive 1 of Participant #1).

## Results

Table 5.2 presents results per individual participant. Due to the low number of participants in this study, statistical tests are not reported, as these were deemed unreliable.

Participants drove on average about 32 km per drive, at a mean speed of 86 km/h (Table 5.2), which is well below the speed limit of 112 km/h (the speed limit on British motorways is 70 mph, which equals 112 km/h). The difference in duration between Drive 1 and Drive 2 is caused by the fact that Drive 1 was about 4 km longer than Drive 2. This was due to the respective entry- and exit points being in different locations.

The 9 participants together manually took control of the automated driving system 9 times, of which 4 were due to the Autopilot failing to anticipate on traffic merging between the LV and PV, 2 to the Autopilot following the undesired line at an exit point, 1 to an unexpected disengagement of the Autopilot, and 1 to the participant disengaging the Autopilot without apparent reason. The remaining Autopilot disengagement occurred for unknown reasons. Furthermore, lane changes were performed 29 and 41 times during the no task (NT) and detection task (DT) condition, respectively, of which 16 and 17 were manual.

Table 5.3 shows the results of the self-report questionnaires. Averaged across the two drives, the mean self-reported overall workload was 19% (21% in Drive 1, 16% in Drive 2; 18% for NT drives, 19% for DT drives). The mean (*SD*) per TLX item was 26% (19%) for Physical Demand, 8% (6%) for Mental Demand, 12% (9%) for Temporal Demand, 27% (29%) for Performance, 17% (13%) for Effort, and 21% (23%) for Frustration. The DSSQ results showed that participants exhibited an overall disengagement from the task, and a worrisome attitude towards the task compared to the pre-task DSSQ (i.e., the standardized change scores are smaller than 0).

Table 5.4 shows the results of the physiological measures. An acclimatization effect can be seen, with the heart rate being lower in Drive 2 than in Drive 1 for 8 out of 9 participants. The SDNN exhibits a negative correlation with the heart rate (see also Heikoo et al., 2017). Here, for 7 of 9 participants, SDNN was higher in Drive 2 than in Drive 1. The LF/HF ratio and respiratory rate remained relatively constant throughout the two drives.

Table 5.2

*Descriptive results (time of day, duration, mean speed, manual overrides, manual and automated lane changes) per participant and drive number.*

PP	Time of day		Duration (s)		Mean speed (km/h)		Manual overrides		Manual lane changes		Automated lane changes	
	Drive 1	Drive 2	Drive 1	Drive 2	Drive 1	Drive 2	Drive 1	Drive 2	Drive 1	Drive 2	Drive 1	Drive 2
	1	2	1	2	1	2	1	2	1	2	1	2
1 (NT, DT)	10:52	11:55	2158	1925	80.5	82.9	0	0	3	1	0	2
2 (DT, NT)	14:29	15:18	2032	1892	85.2	84.1	0	0	1	4	5	1
3 (NT, DT)	10:43	11:46	2057	1653	85.4	96.5	1	1	2	4	1	4
4 (DT, NT)	14:40	15:20	1440	1405	85.6	86.1	0	0	3	0	2	2
5 (NT, DT)	10:33	11:28	2073	1871	83.6	85.1	2	0	0	1	3	2
6 (DT, NT)	14:41	15:33	2013	1834	86.1	87.0	0	0	2	2	1	1
7 (NT, DT)	10:25	11:18	1991	1813	87.1	87.9	0	1	1	2	2	3
8 (DT, NT)	14:34	15:34	1928	1800	88.9	88.2	2	0	2	1	1	2
9 (NT, DT)	10:40	11:42	2008	1886	86.1	84.6	1	1	3	1	1	4
Average			2033	1834	85.4	86.9	0.67	0.33	1.89	1.78	1.78	2.33

Note. NT = No Task, DT = Detection Task. Participant #4 did not complete the entire route because the batteries of the car were emptying and the car needed to be charged. This participant was excluded from the calculation of the average duration.

Table 5.3

*Self-reported overall workload (TLX OW) and standardised change scores of self-reported stress (DSSQ) per participant and drive number.*

PP	TLX OW (%)		DSSQ engagement		DSSQ distress		DSSQ worry	
	Drive 1	Drive 2	Drive 1	Drive 2	Drive 1	Drive 2	Drive 1	Drive 2
1 (NT, DT)	12	8	-0.10	-0.21	-0.43	-0.49	-0.35	-0.57
2 (DT, NT)	18	14	-0.12	-0.18	-0.97	-1.13	0.17	-0.16
3 (NT, DT)	45	46	-1.10	-0.19	0.67	0.23	-0.78	0.14
4 (DT, NT)	17	3	-0.01	0.26	-0.82	-1.41	-2.34	-1.93
5 (NT, DT)	31	20	-0.37	-0.61	0.40	-0.16	-0.51	-0.72
6 (DT, NT)	14	14	0.18	0.46	-0.30	-0.63	-0.34	-0.32
7 (NT, DT)	9	6	-1.48	-0.52	0.46	0.26	0.49	-0.68
8 (DT, NT)	40	25	-2.60	-1.88	2.67	-0.90	0.81	1.59
9 (NT, DT)	5	7	-1.43	-0.02	1.69	1.13	-1.13	-1.69
Average	21	16	-0.78	-0.32	0.38	-0.34	-0.44	-0.48

Table 5.4

*Cardiovascular and respiratory results per participant and drive number.*

PP	Heart rate (bpm)		SDNN (ms)		LF/HF ratio		Respiratory rate (bpm)	
	Drive 1	Drive 2	Drive 1	Drive 2	Drive 1	Drive 2	Drive 1	Drive 2
1 (NT, DT)	61.3	61.9	44.7	46.2	0.79	0.72	15.3	15.7
2 (DT, NT)	99.4	90.4	20.3	33.6	1.52	2.00	N.A.	N.A.
3 (NT, DT)	59.7	56.6	50.2	53.9	0.97	0.73	14.1	15.1
4 (DT, NT)	72.2	67.5	48.3	59.4	1.02	1.14	18.5	17.8
5 (NT, DT)	69.0	66.4	55.7	47.7	1.26	1.26	15.6	14.9
6 (DT, NT)	82.1	73.0	43.6	42.4	0.91	0.91	18.4	17.2
7 (NT, DT)	59.2	58.9	44.9	51.5	0.89	0.94	19.2	19.7
8 (DT, NT)	80.4	70.7	35.3	39.5	0.99	0.96	20.2	20.3
9 (NT, DT)	56.6	52.2	93.2	101.3	1.18	1.02	18.4	18.7
Average	71.1	66.4	48.5	52.8	1.06	1.08	17.0	16.5

Note. The respiratory rate of Participant #2 is not provided, as there was no clear peak value to be identified from the Fourier transformation.

Due to technical issues, participants #3, 4, and 6 had no button press data for 14, 16, and 10 bridges respectively. Hit rates and false alarm rates were calculated for the remaining number of bridges for these participants. The mean (*SD*) hit and false alarm rates of the bridges were 88.0% (16.0%) and 0.9% (1.5%), respectively. The lowest hit rate was 47.1%, whereas two participants had hit rates of 100% (Figure 5.7). The mean hit rate and mean false alarm rate correspond to a perceptual sensitivity ( $d'$ ) of 3.52 and a response bias ( $\beta$ ) of 7.89 (Stanislaw & Todorov, 1999). These results indicate that participants were well able to distinguish the bridges from the non-bridges with a conservative response strategy.

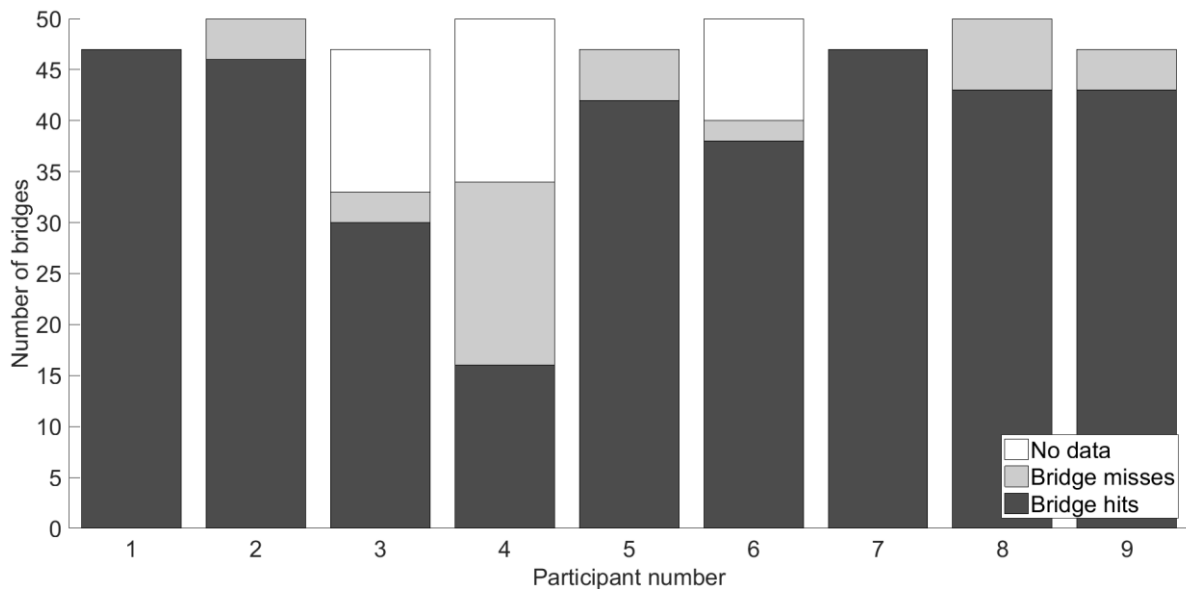


Figure 5.7. Number of hits and misses per participant. The number of false alarms was 1, 0, 0, 0, 1, 0, 0, 0, 2 for participants 1–9.

## Discussion

This study aimed to measure levels of workload and stress during automated driving with the Tesla Autopilot. The literature has shown that automated driving yields low ratings of self-reported overall workload (averaging at 23%, see De Winter et al., 2014). A previous study using the Tesla Autopilot has found overall workload scores ranging from 10% for experienced Tesla drivers on an empty highway to 43% on a city ring with drivers who had not driven in the Tesla before (Table 5.1; Stapel et al., 2017). Another on-road experiment

found high workload for automated driving compared to manual driving, with overall workload scores for automated driving being 42% (Banks & Stanton, 2016). It was unclear whether the novelty of the automation in Banks and Stanton (2016) created elevated levels of workload, so they proposed extended exposure to automation, which was the purpose of the current study.

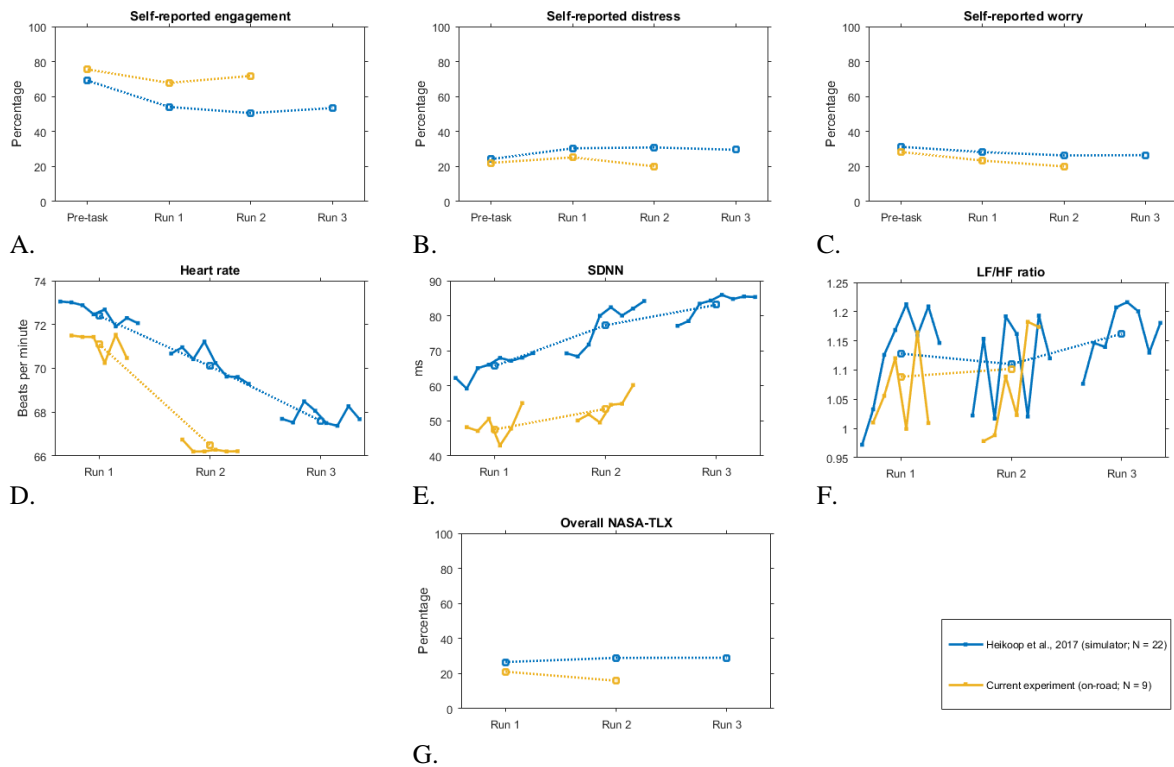
Our results of a 2 x 32 min of automated driving showed that the mean overall workload dropped from 21% in Drive 1 to 16% in Drive 2. In other words, automated driving involves a level of self-reported workload that is within the range of the workload observed in driving simulators (see Table 5.1, which shows a minimum overall workload of 11% and a maximum of 36%). The fact that the participants had to follow a lead vehicle may have contributed to the low overall workload by limiting their decision making requirements. The workload was particularly low for the Physical Demand item, which may be because the participants did not have to move the pedals or steering wheel for large portions of the time. Similarly, the average heart rate and respiratory rate were close to the resting rates of a typical person (American Heart Association, 2015; Lindh, Pooler, Tamparo, & Dahl, 2009). Thus, automated driving in the present experiment could not be considered demanding or stressful.

Compared to the detection task in Heikoop et al. (2017), participants in the current study performed somewhat worse (hit rate of 95% in Heikoop et al.; 88% in the present study). Furthermore, self-reported workload in the present study was not substantially different between the DT and NT conditions. It is possible that, for safety reasons, participants in the present study were trying harder to stay alert for the primary driving task as compared to the participants in the simulator study, thereby having a lower incentive to detect the bridges. Indeed, with the Tesla Autopilot, participants do have to remain alert due to the potential need to intervene and take manual control of the vehicle. In a survey study by Dikmen and Burns (2016), 62% of Autopilot users reported that they had experienced at least

one unexpected or unusual behaviour when driving in automated mode. In our experiment, the 9 participants took over manual control 9 times, for a variety of reasons, and performed 70 lane changes, of which 47% were manual. The relatively large percentage of manual lane changes may be due to the somewhat cumbersome technique required to perform an automated lane change. For an automated lane change to succeed, the driver has to press the indicator stalk while having his/her hands on the steering wheel with enough weight for the Autopilot to recognize their presence. In our experiment, this often resulted in a slight turn of the steering wheel, disengaging the Autopilot, after which the lane change had to be performed manually.

Comparing participants' self-reported engagement during a previous driving simulator study featuring similar methods (Heikoop et al., 2017) with the current on-road study (Figure 5.8), it can be seen that participants felt relatively engaged during the on-road study. The relatively high level of engagement may be because the participants in the present study prioritized safety and tried to stay alert, as noted above. Participants of this on-road study reported relatively similar levels of distress and worry compared to the participants in the simulator study. This may be because the participants in the on-road study were professional drivers who were used to driving with advanced driver assistance systems. Therefore, even though they maintained higher engagement levels than the participants in the simulator study, this did not translate into more distress or worry. Furthermore, the overall self-reported workload was low (a mean score of 19% on a scale from 0% to 100%) despite the fact that participants were exposed to the Tesla for the first time. Both the on-road and simulator-based studies found that self-reported workload remained approximately constant, heart rate decreased, and SDNN increased as a function of time. Previous on-road studies into manual driving also found that the heart rate tends to decrease with time on task (Lisper, Laurell, &

Stening, 1973; Schmidt, Schrauf, Simon, Fritzsche, Buchner, & Kincses, 2009). Finally, the LF/HF ratio in the simulator and on the road were within the same range (Fig. 8).



*Figure 5.8.* A–C) DSSQ results, D–F) Cardiovascular measures, and G) Self-reported workload, for two experiments (Exp. 1: Heikoop et al., 2017 [N = 22], Exp. 2: current experiment [N = 9]). The individual drives of the two experiments lasted 40 and 32 min, respectively. The DSSQ percentages were calculated with respect to the minimum and maximum scores achievable per DSSQ scale.

## Conclusions and Recommendations

This experiment complements existing research on stress and workload during automated driving (cf. Table 5.1), and may form a basis for more extensive research on this topic. The results point to an effect of acclimatization as demonstrated by a drop of perceived workload over time, and a decrease in heart rate.

Our sample was small ( $N = 9$ ); in order to acquire greater statistical power, replication studies with more participants are advised. Nonetheless, our study produced insights into the effects of workload and acclimatization to automation, and could serve as a foundation for future research into this phenomenon. This experiment used expert drivers, and it remains to be investigated how the results translate to less trained drivers. It is likely that the participants



in this study, who were experienced with various supervars and had completed advanced driver training, are more adept to novel technology and less likely to be stressed than the general population. Stapel et al. (2017) found in their on-road study that people who have experience with driving in a Tesla reported lower workload than people who had not driven a Tesla before, both during manual and automated driving.

The methods used for psychophysiological measurement need consideration. Although the LF/HF ratio is regarded as a valid measure of workload (Cinaz et al., 2013; Suriya-Prakash et al., 2015), the results of the LF/HF ratio in this experiment did not show the same time-on-task effects as the self-reported overall workload. The lack of sensitivity of the LF/HF ratio could be due to various confounding effects such as driver posture, vibrations in the vehicle, or a dependency on the heart rate itself (as also discussed by Heikoop et al., 2017).

The focus in this experiment was on how drivers are affected by automated driving over time. Future on-road research could include a control condition in which people drive manually (cf. Stapel et al., 2017) or include different levels of automation such as driving with ACC, or driving with ACC and steer assist (see Naujoks, Purucker, & Neukum, 2016). In our experiment, control transitions were not wanted, yet occurred several times per participant. A closed-track on-road experiment (cf. Albert, Lange, Schmidt, Wimmer, & Bengler, 2015; Omae, Hashimoto, Sugamoto, & Shimizu, 2005) could investigate the psychological effects of transition of control to and from automated driving in a controlled manner. Furthermore, it could be investigated whether the present findings generalize to situations with actual hazards for which manual intervention is necessary. It is possible that drivers would score better in an environment in which the target stimuli represent actual safety-critical events. Finally, it remains to be studied how our findings generalize to other driving scenarios, such as driving during rush hours.

## References

- Albert, M., Lange, A., Schmidt, A., Wimmer, M., & Bengler, K. (2015). Automated driving – Assessment of interaction concepts under real driving conditions. *Proceedings of the 6<sup>th</sup> International Conference on Applied Human Factors and Ergonomics (AHFE2015), and the Affiliated Conferences*, 2832-2839. doi: 10.1016/j.promfg.2015.07.767
- American Heart Association. (2015). Target heart rates. January 2015, from [http://www.heart.org/HEARTORG/HealthyLiving/PhysicalActivity/FitnessBasics/Target-Heart-Rates\\_UCM\\_434341\\_Article.jsp#.V8Q9T\\_krKWg](http://www.heart.org/HEARTORG/HealthyLiving/PhysicalActivity/FitnessBasics/Target-Heart-Rates_UCM_434341_Article.jsp#.V8Q9T_krKWg)
- Banks, V. A., & Stanton, N. A. (2016). Keep the driver in control: Automating automobiles of the future. *Applied Ergonomics*, 53, 389-395. doi: 10.1016/j.apergo.2015.06.020
- Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F*, 35, 75-84. doi: 10.1016/j.trf.2015.10.005
- Borojeni, S. S., Chuang, L., Heuten, W., & Boll, S. (2016). Assisting drivers with ambient take-over requests in Highly Automated Driving. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '16)*, 237-244. doi: 10.1145/3003715.3005409
- Brackstone, M., & McDonald, M. (2007). Driver headway: How close is too close on a motorway? *Ergonomics*, 50, 1183-1195. doi: 10.1080/00140130701318665
- Casner, S. M., Hutchins, E. L., & Norman, D. (2016). The challenges of partially automated driving. *Communications of the ACM*, 59, 70–77. doi: 10.1145/2830565
- Cinaz, B., Arnrich, B., La Marca, R., & Tröster, G. (2013). Monitoring of mental workload levels during an everyday life office-work scenario. *Personal and Ubiquitous Computing*, 17, 229–239. doi: 10.1007/s00779-011-0466-1
- Damböck, D., Weißgerber, T., Kienle, M., & Bengler, K. (2013). Requirements for cooperative vehicle guidance. *Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems*, 1656-1661. doi: 10.1109/itsc.2013.6728467
- Davis, J., Animashaun, A., Schoenherr, E., & McDowell, K. (2008). Evaluation of semi-autonomous convoy driving. *Journal of Field Robotics*, 25, 880–897. doi: 10.1002/rob.20263
- De Waard, D. (1996). *The measurement of drivers mental workload*. (Doctoral dissertation), Rijksuniversiteit Groningen.
- De Winter, J. C. F., Happee, R., Martens, M., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, 27, 196–217. doi: 10.1016/j.trf.2014.06.016
- De Winter, J. C. F., Stanton, N. A., Price, J. S., & Mistry, H. (2016). The effects of driving with different levels of unreliable automation on self-reported workload and secondary task performance. *International Journal of Vehicle Design*, 70, 297–324. doi: 10.1504/IJVD.2016.076736
- Dikmen, M., & Burns, C. M. (2016). Autonomous driving in the real world: Experiences with Tesla Autopilot and Summon. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '16)*, 225-228. doi: 10.1145/3003715.3005465
- Endsley, M. R. (2017). Autonomous driving systems: A preliminary naturalistic study of the Tesla Model S. *Journal of Cognitive Engineering and Decision Making*. doi: 10.1177/1555343417695197

- Eriksson, A., Banks, V. A., & Stanton, N. A. (2017). Transition to manual: Comparing simulator with on-road control transitions. *Accident Analysis & Prevention*, 102, 227-234. doi: 10.1016/j.aap.2017.03.011
- Eriksson, A., & Stanton, N. A. (2017). Take-over time in highly automated vehicles: Non-critical transitions to and from manual control. *Human Factors*. doi: 10.1177/0018720816685832
- Fairclough, S. H., & Venables, L. (2005). Prediction of subjective states from psychophysiology: A multivariate approach. *Biological Psychology*, 71, 100-110. doi: 10.1016/biopsycho.2005.03.007
- Fallahi, M., Motamedzade, M., Heidarimoghadam, R., Soltanian, A. R., & Miyake, S. (2016). Effects of mental workload on physiological and subjective responses during traffic density monitoring: A field study. *Applied Ergonomics*, 52, 95-103. doi: 10.1016/j.apergo.2015.07.009
- Full self-driving hardware on all cars. (2015). from [https://www.tesla.com/en\\_GB/autopilot](https://www.tesla.com/en_GB/autopilot)
- Hancock, P. A. (2015). Automobility: The coming use of fully-automated on-road vehicles. *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 137-139. doi: 10.1109/COGSIMA.2015.7108188
- Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, 50, 904-908. doi: 10.1177/154193120605000909
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139-183. doi: 10.1016/S0166-4115(08)62386-9
- Haslbeck, A., & Hoermann, H.-J. (2016). Flying the needles: Flight deck automation erodes fine-motor flying skills among airline pilots. *Human Factors*, 58, 533-545. doi: 10.1177/0018720816640394
- Healey, J. A., & Picard, R. W. (2004). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6, 156-166. doi: 10.1109/TITS.2005.848368
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2017). Effects of platooning on signal-detection performance: A driving simulator study. *Applied Ergonomics*, 60, 116-127. doi: 10.1016/j.apergo.2016.10.016
- Helton, W. S., Warm, J. S., Matthews, G. S., Corcoran, K. J., & Dember, W. N. (2002). Further tests of an abbreviated vigilance task: Effect of signal salience and jet aircraft noise on performance and stress. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46, 1546-1550. doi: 10.1177/154193120204601704
- Jorna, P. G. A. M. (1992). Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. *Biological Psychology*, 34, 237-257. doi: 10.1016/0301-0511(92)90017-O
- Kaber, D. K., & Endsley, M. R. (1997). Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. *Process Safety Progress*, 16, 126-131. doi: 10.1002/prs.680160304
- Large, D. R., Banks, V. A., Burnett, G. E., Baverstock, S., & Skrypchuk, L. (2017). Exploring the behaviour of distracted drivers during different levels of automation in driving. *Proceedings of the 5th International Conference on Driver Distraction and Inattention (DDI2017)* (pp. 1-18). Paris, France.
- Lindh, W. Q., Pooler, M. S., Tamparo, C. D., & Dahl, B. M. (2009). Vital signs and measurements. In W. Q. Lindh, M. S. Pooler, C. D. Tamparo & B. M. Dahl (Eds.),

- Delmar's comprehensive medical assisting: Administrative and clinical competencies.* (pp. 573). Clifton Park, NY: Cengage Learning Inc. .
- Lisper, H.-O., Laurell, H., & Stening, G. (1973). Effects of experience of the driver on heart-rate, respiration-rate, and subsidiary reaction time in a three hours continuous driving task. *Ergonomics*, 16, 501-506. doi: 10.1080/0014137308924539
- Mackworth, J. F. (1964). Performance decrement in vigilance, threshold, and high-speed perceptual motor tasks. *Canadian Journal of Psychology*, 18, 209-223. doi: 10.1037/h0083302
- Manawadu, U., Ishikawa, M., Kamezaki, M., & Sugano, S. (2015). Analysis of individual driving experience in autonomous and human-driven vehicles using a driving simulator. *Proceedings of the IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, 299-304. doi: 10.1109/AIM.2015.7222548
- Matthews, G. S. (2014). *Task-induced fatigue, operator engagement and the control of attention*. PowerPoint slides.
- Matthews, G. S., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., . . . Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, 2, 315-340. doi: 10.1037/1528-3542.2.4.315
- Matthews, G. S., Campbell, S.E., & Falconer, S. (2000). Personality, energetics and individual differences in motivation. *Proceedings of the Tenth European Conference on Personality*. Krakow, Poland.
- Matthews, G. S., Szalma, J., Panganiban, A. R., Neubauer, C., & Warm, J. S. (2013). Profiling task stress with the Dundee Stress State Questionnaire. In L. Cavacanti & S. Azevedo (Eds.), *Psychology of stress: New research* (pp. 49-91). Hauppauge, NY: Nova Science Publishers, Inc.
- McDowell, K., Nunez, P., Hutchins, S., & Metcalfe, J. S. (2008). Secure mobility and the autonomous driver. *IEEE Transactions on robotics*, 24, 688-697. doi: 10.1109/TRO.2008.924261
- McKenzie, B., & Rapino, M. (2011). Commuting in the United States: 2009. *US Department of Commerce, Economics and Statistics Administration, US Census Bureau*.
- Metzger, U., & Parasuraman, R. (2001). The role of the air traffic controller in future air traffic management: An empirical study of active control versus passive monitoring. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43, 519-528. doi: 10.1518/001872001775870421
- 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. *Transportation Research Part F: Traffic Psychology and Behaviour*, 38, 67-82. doi: 10.1016/j.trf.2016.01.011
- Neubert, L., Santen, L., Schadschneider, A., & Schreckenberg, M. (1999). Single-vehicle data of highway traffic - A statistical analysis. *Physical Review E*, 60, 6480-6490. doi: 10.1103/PhysRevE.60.6480
- NHTSA. (2017). Automatic vehicle control systems. Investigation PE 16-007. from <https://static.nhtsa.gov/odi/inv/2016/INCLA-PE16007-7876.pdf>
- Omae, M., Hashimoto, N., Sugamoto, T., & Shimizu, H. (2005). Measurement of driver's reaction time to failure of steering controller during automated driving *Review of Automotive Engineering*, 26, 213-215.
- Pasaoglu, G., Fiorello, D., Martino, A., Zani, L., Zubaryeva, A., & Thiel, C. (2014). Travel patterns and the potential use of electric cars—Results from a direct survey in six European countries. *Technological Forecasting and Social Change*, 87, 51-59. doi: 10.1016/j.techfore.2013.10.018

- Petermeijer, S. M., Bazilinskyy, P., & De Winter, J. C. F. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied Ergonomics*, 62, 204-215. doi: 10.1016/j.apergo.2017.02.023
- Petermeijer, S. M., Cieler, S., & De Winter, J. C. F. (2017). Comparing spatially static and dynamic vibrotactile take-over requests in the driver seat. *Accident Analysis and Prevention*, 99, 218-227. doi: 10.1016/j.aap.2016.12.001
- Saxby, D. J., Matthews, G. S., 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. *Journal of Experimental Psychology: Applied*, 19, 287-300. doi: 10.1037/a0034386
- Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis and Prevention*, 41, 1087-1093. doi: 10.1016/j.aap.2009.06.007
- Schwalk, M., Kalogerakis, N., & Maier, T. (2015). Driver support by a vibrotactile seat matrix - Recognition, adequacy and workload of tactile patterns in take-over scenarios during automated driving. *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics (AHFE2015), and the Affiliated Conferences*, 2466-2473. doi: 10.1016/j.promfg.2015.07.507
- Shaffer, F., & Combatalade, D. C. (2013). Don't add or miss a beat: A guide to cleaner Heart Rate Variability recordings. *Biofeedback*, 41, 121-130. doi: 10.5298/1081-5937-41.3.04
- Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. MIT press.
- Song, M., & Wang, J-H. (2010). Studying the tailgating issue in Rhode Island and its treatment. *Proceedings of the 51st Annual Transportation Research Forum*. Arlington, VA.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31, 137-149. doi: 10.3758/BF03207704
- Stanton, N. A., & Marsden, P. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science*, 24, 35-49. doi: 10.1016/S0925-7535(96)00067-7
- Stanton, N. A., Young, M., & McCaulder, B. (1997). Drive-by-wire: The case of mental workload and the ability of the driver to reclaim control. *Safety Science*, 27, 149-159. doi: 10.1016/S0925-7535(97)00054-4
- Stapel, J., Mullakka-Babu, F. A., & Happee, R. (2017). *Driver behavior and workload in an on-road automated vehicle*. Paper presented at the Road Safety & Simulation International Conference, The Hague, The Netherlands.
- Suriya-Prakash, M., John-Preetham, G., & Sharma, R. (2015). Is heart rate variability related to cognitive performance in visuospatial working memory? *PeerJ PrePrints*, 3, e1711. doi: 10.7287/peerj.preprints.1377v1
- Tanner Jr., W. P., Wilson, P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61, 401-409. doi: 10.1037/h0058700
- Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: Standards of measurement, physiological interpretation and clinical use. *Circulation*, 93, 1043-1065. doi: 10.1161/01.CIR.93.5.1043
- Teichner, W. H. (1974). The detection of a simple visual signal as a function of time of watch. *Human Factors*, 16, 339-353. doi: 10.1177/001872087401600402

- Treiber, M., Kesting, A., & Helbing, D. (2006). Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps. *Physical Review E*, 74, 1-10. doi: 10.1103/PhysRevE.74.016123
- Vidulich, M. A., & Tsang, P. S. (2012). Mental workload and situation awareness. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (4th ed.): John Wiley & Sons, Inc.
- Vollmer, M. (2015). A robust, simple and reliable measure of heart rate variability using relative RR intervals. *IEEE Conference on Computing in Cardiology Conference (CinC)*, 609-612. doi: 10.1109/CIC.2015.7410984
- Young, M. S. (2000). *Attention, automaticity, and automation: New perspectives on mental underload and performance*. (Doctoral dissertation), University of Southampton, UK.
- Young, M. S., & Stanton, N. A. (2004). Taking the load off: Investigations of how adaptive cruise control affects mental workload. *Ergonomics*, 47, 1014-1035. doi: 10.1080/00140130410001686348
- Young, M. S., & Stanton, N. A. (2007). What's skill got to do with it? Vehicle automation and driver mental workload. *Ergonomics*, 50, 1324-1339. doi: 10.1080/00140130701318855

## Supplementary materials

Photos of the bridges obtained with Google Streetview.

<https://www.dropbox.com/sh/bln7zbs0o1h9af7/AAAX2SFO2Z9vEfqu-dX4Wpiba?dl=0>

# 6

## Discussion and Conclusions

---

## Conclusions of the Thesis

This thesis aimed to investigate the effects of prolonged automated platooning on drivers' psychological state. By means of a data-driven literature study (Chapter 2), key psychological constructs were identified as well as relationships between these constructs. Accordingly, Chapter 2 presented a revision of a model on psychological constructs in automated driving initially proposed by Stanton and Young (2000) (Figure 6.1). Three subsequent experiments assessed several type-of-task effects as well as time-on-task effects on a selection of psychological constructs during extended periods of driving in a platoon (Chapters 3 & 4) or car following on a highway (Chapter 5).

Specifically, Chapter 3 investigated the effects of a visual detection task, a voluntary task, and no task on driver workload and stress. Chapter 4 assessed the effects of mental task demands on situation awareness, and the development of mental models over time. Finally, Chapter 5 aimed to investigate the generalizability of the results found in Chapter 3 in an on-road experiment.

The following conclusions are drawn from this thesis:

- 1) According to the literature in the domain of Human Factors of automated driving, several causal relationships exist between psychological constructs, as depicted in Figure 6.1 (Chapter 2).
- 2) The notion that 'man is a poor monitor' does not apply to our platooning task. That is, drivers are able to maintain a high percentage of correct responses in a visual detection task during simulator-based platooning (Chapter 3).
- 3) During an automated platooning drive, having the freedom to do whatever one wants (i.e., performing the voluntary task condition) is the least stressful and demanding for drivers. This was evidenced by the relatively low LF/HF ratio, and the relatively low self-reported workload, distress, and worry, as compared to having nothing to do



except to watch the road (i.e., performing the no-task condition) and compared to performing the visual detection task (Chapter 3).

- 4) Certain effects observed in the simulator generalize to a real-world environment. Specifically, a drop in heart rate over time and self-reported workload was found in both the simulator study and the on-road study, which can be seen as an indication towards acclimatization to the automation (Chapter 3 & 5). Furthermore, the absolute levels of self-reported workload, engagement, distress, worry, and cardiovascular measures are in approximate agreement between the simulated and on-road environments (Chapter 5).
- 5) High mental demands (induced using a 2-back letter task) reduce drivers' situation awareness as compared to low task demands (Chapter 4).
- 6) Drivers in an automated platoon generate increasingly complex mental models as a function of the amount of time spent driving in the platoon (Chapter 4).

When interpreting the results of the three experiments together, the overall conclusion is that drivers of a platoon are able to remain vigilant for at least 30 to 40 min during one automated platooning drive. At the same time, it is clear that drivers should not be expected to remain *perfectly* vigilant, as the hit rates were 95% (Chapter 3) and 88% (Chapter 5), not 100%. A hit rate as high as 95% (i.e., miss rate of 5%) may be insufficient in on-road driving, because if a miss occurs during a critical situation, this may have serious consequences. Furthermore, the requirement to remain vigilant does appear to cause some mental demands/stress (see Conclusion 3 above). In fact, drivers are inclined to perform irrelevant tasks (such as interacting with one's mobile phone or enjoying the scenery) when told this is permitted, despite being instructed to avoid accidents and to intervene whenever a critical situation appears (Chapter 3; see also Llaneras, Salinger, & Green, 2013; Omae, Hashimoto,

Sugamoto, & Shimizu, 2005). Moreover, even if drivers are watching the road ahead, this does not guarantee they actually see the road ahead, as any non-driving task that adds mental demands (on top of their regular driving demands) interferes with situation awareness (Chapter 4; see also Ma & Kaber, 2005; Rogers, Zhang, Kaber, & Liang, 2011).

### **Reflection on the Proposed Psychological Model**

The psychological model proposed in Chapter 2 (Figure 6.1) provides confirmatory consensus as well as new insights regarding psychological constructs that play a role in automated driving. In the first place, the model establishes links between task demands, mental workload, and attention, which are amongst the most commonly mentioned links in the literature. Indeed, in the field of automated driving, the importance of appropriate levels of attention, and having an optimal level of workload (i.e., no under- or overload) has been stressed since the mid-1990s (Stanton & Marsden, 1996). Also, we found that the constructs that have overlapping symptoms (e.g., related to mood affect), yet have distinctly different effects (e.g., on performance), namely fatigue and stress, are commonly used in the literature (e.g., Desmond & Matthews, 2009; Hockey, 1997; Matthews, 2002). Concurrently, our model provides new insights regarding the fact that mediating constructs are often overlooked in published research. For instance, the causal relationship from mental workload to situation awareness is mediated through attention (see also Chapter 4 and e.g., Bellet, Bailly-Asuni, Mayenobe, & Banet, 2009; Wickens, 2002).

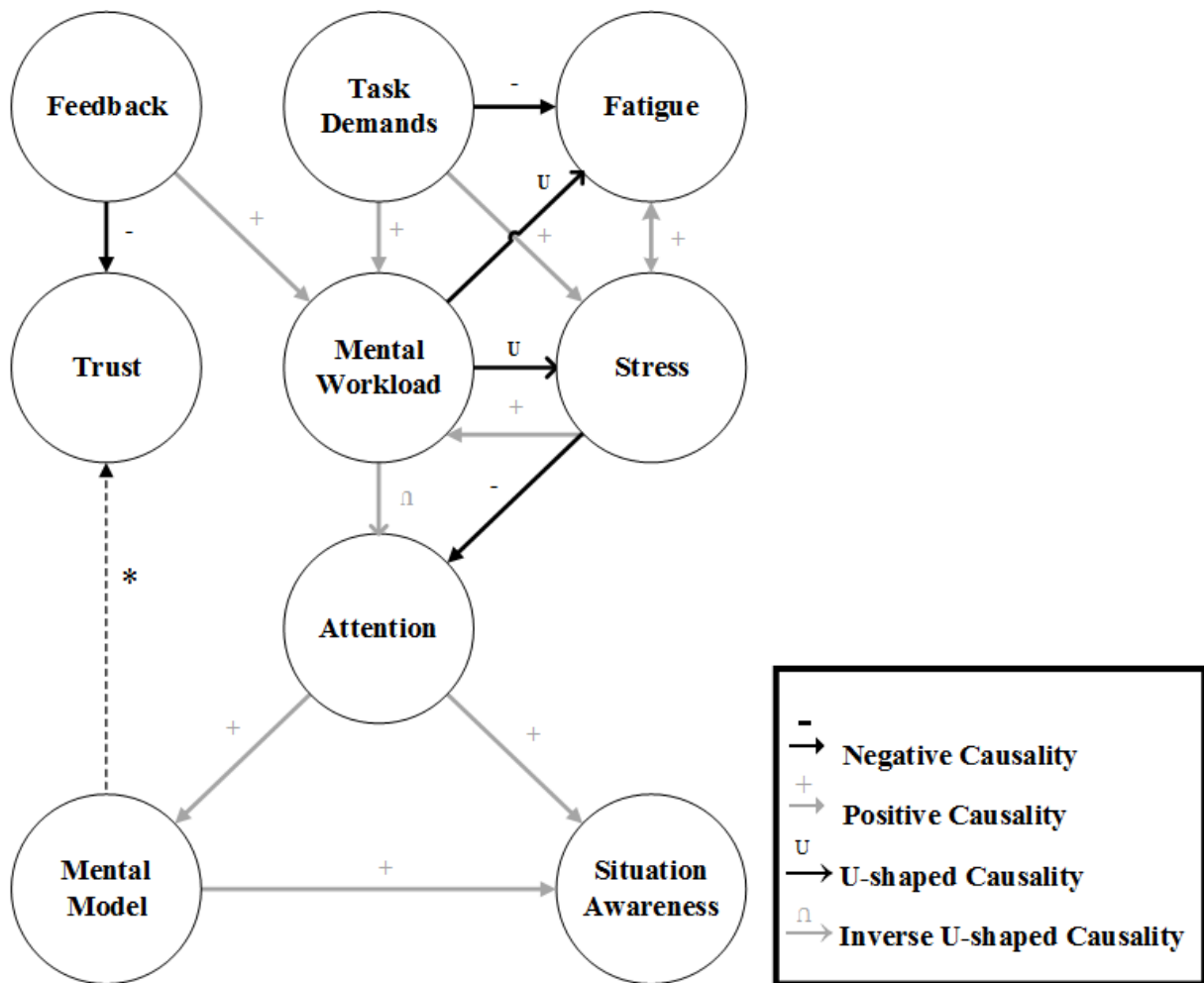


Figure 6.1. The proposed psychological model (Chapter 2).

In conclusion, the proposed model serves as an indication of what the consensus is in the automated driving domain regarding psychological constructs and their interactions (see Chapter 2 for a more thorough discussion). Next, in an attempt to validate the proposed model, it was considered to test the model empirically. An initial step towards model testing has been performed in this thesis, by assessing specific sections of the model.

## The Issue of Construct Proliferation

An observation made in Chapter 2 is that the domain of Human Factors suffers from construct proliferation. For example, the psychological construct noted here as “mental workload” was connected to more than 20 different descriptions, ranging from “mental

processing” to “cognitive activity” (Table 2.3). It is noted that proliferation occurs certainly not always because researchers coin new constructs on purpose, but often due to context-specific referencing.

The work presented in this thesis also inevitably suffered from construct proliferation, not solely due to the cited literature, but also due to the limited tools at hand used in the experiments, and the lack of discriminatory power of the results (i.e., the limited sample sizes). For instance, it is a well-known fact that heart rate is influenced by several factors, such as respiration and physical activity (Appelhans & Luecken, 2006; Brookhuis & De Waard, 2010; Mehler, 2015; Mulder, 1992), whereas in this thesis we attributed a decrease in heart rate to decreased stress, increased fatigue, as well as acclimatization to the experimental conditions (see e.g., Healey & Picard, 2004; Lal & Craig, 2000). Thus, one may wonder whether any 1-to-1 connection between physiological measures (e.g., heart rate) and psychological constructs (e.g., fatigue) can justifiably be made (Cacioppo & Tassinari, 1990). The bidirectional positive relation between stress and fatigue depicted in the model (Figure 6.1) is the opposite from the relation we attributed it in this thesis. This may be because our model does not specify the types of fatigue. In the experiments performed in this thesis only one specific aspect of fatigue was elicited: passive fatigue (see Desmond & Hancock, 2001) because of the monotone automated platooning scenarios without need for manual intervention. Experiments that would elicit active fatigue (e.g., driving with harsh wind gusts, like in Saxby, Matthews, Hitchcock, & Warm, 2007) could show a different relation between stress and fatigue.

Furthermore, several of the measures presented in this thesis were found to be strongly correlated. Specifically, heart rate correlated strongly with measures of heart rate variability (SDNN & LF/HF ratio, see Table 6.1). Of course, it is expected that physiological measures do have (co)relations, as our model predicts that constructs are related to each other (Figure

6.1). However, correlations between physiological measures (e.g., between heart rate and SDNN) may be purely a physical necessity (i.e., a higher mean NN [i.e., lower mean heart rate] is accompanied with a higher standard deviation of NN) rather than having distinct psychological causes. Thus, whether SDNN uniquely measures mental workload, or whether it is merely a by-product of heart rate, cannot be unambiguously established.

To measure fatigue and stress, we used the Dundee Stress State Questionnaire (DSSQ), which distinguishes between as many as 13 facets of stress (Chapter 3). Although these 13 facets carry face validity and have an empirical basis (e.g., via factor analyses with large sample sizes; Matthews et al., 1999), we decided to proceed with higher-order factors of the DSSQ (distress, engagement, worry). However, these higher-order factors may be subject to even higher order factors. To illustrate, Table 6.1 shows the Spearman correlation matrix of selected measures from Experiments 1, 2, and 3 combined, where it can be seen, for example, that TLX Frustration correlated relatively strongly ( $\rho = .47$ ) with DSSQ Distress (i.e., more strongly than with the other five TLX items). Indeed, Matthews et al. (1999) found similar results ( $r = .58$ ). This relatively strong correlation may signal that frustration and distress reflect the same higher-order construct. Similarly, it needs consideration whether the constructs presented in the model (Figure 6.1) are unique enough to be regarded as empirically distinguishable.

In conclusion, this section aims to make a call for general definitions of (higher-level) psychological constructs, much like the proposition made by Musek (2007), who showed that it is feasible to merge the Big Five personality factors into one general personality *trait*. This ‘Big One’ in turn could improve understanding of personality, as it could define the basic root construct of human personality, which then optionally could be subdivided into more fine-grained traits. Since several psychological *state* constructs are known to correlate heavily,

perhaps a definition of one encompassing construct (e.g., a general factor of arousal) could similarly improve understanding of psychology.

Finally, it is emphasized that consensus in the use of psychological constructs must be reached in order to be able to understand, and avoid unnecessarily replicating, each other's research (see e.g., Le, Schmidt, Harter, & Lauver, 2010; Shaffer, DeGeest, & Li, 2017, who provide a similar suggestion). Indeed, the importance of model parsimony is well recognized in science and dates back to Aristotle's *Posterior Analytics* (350 BC), in which he writes: "We may assume the superiority [*ceteris paribus*] of the demonstration which derives from fewer postulates or hypotheses-in short from fewer premises" (translated by Mure, 2007).

### **Simulator versus On-Road: A Comparison of Results**

Chapter 5 presented an on-road experiment that aimed to assess the same psychophysiological and self-report measures as used in Chapter 3, to identify differences and similarities with simulator experimentation. Driving simulators are beneficial to researchers, as they are relatively easy to use, enable reproduction of scenarios, and provide a safe environment (Reed & Green, 1999). However, concerns exist about their fidelity (De Winter, Van Leeuwen, & Happee, 2012). One consideration for the use of on-road experimentation is the lack of physical crash risk in simulators, which could potentially lead to low fear and complacent behaviour (cf. Hallvig et al., 2013). On the other hand, for research purposes, it may prove difficult to acquire ethics approval, or a representative pool of participants, as the on-road environment is less controllable, and therefore participants may be required to have the appropriate driving experience and licensing.

Table 6.1

*Spearman rank-order correlations among personal characteristics and driver-state measures (N = 64).*

	Age	Gender	Driving experience	Driving frequency	Mileage	Heart rate	SDNN	LF/HF ratio	DSSQ Engagement	DSSQ Distress	DSSQ Worry	TLX – Mental	TLX – Physical	TLX – Temporal	TLX – Performance	TLX – Effort	TLX – Frustration	TLX – Overall
Age																		
Gender	-0.17																	
Driving experience	0.89	-0.10																
Driving frequency	0.47	-0.26	0.51															
Mileage	0.42	-0.27	0.44	0.80														
Heart rate	-0.15	0.04	-0.20	0.01	-0.09													
SDNN	-0.32	0.00	-0.15	-0.24	-0.08	-0.63												
LF/HF ratio	0.14	0.27	0.13	0.15	0.05	0.45	-0.49											
DSSQ Engagement	0.00	0.06	-0.11	0.04	-0.14	0.24	-0.37	0.16										
DSSQ Distress	0.02	-0.11	0.12	-0.07	0.16	-0.20	0.17	-0.08	-0.34									
DSSQ Worry	0.00	-0.18	-0.02	0.02	0.11	-0.04	-0.09	-0.05	-0.08	0.06								
TLX – Mental	0.02	0.02	0.13	-0.07	-0.11	0.18	0.03	0.05	0.16	0.10	-0.02							
TLX – Physical	0.16	-0.14	0.23	0.11	-0.06	0.12	-0.08	-0.03	0.02	0.21	0.18	0.36						
TLX – Temporal	0.12	-0.24	0.16	0.10	0.04	0.18	-0.20	-0.02	0.13	0.10	0.25	0.45	0.35					
TLX – Performance	-0.02	-0.09	-0.05	-0.09	-0.08	-0.03	-0.09	-0.01	-0.06	0.34	0.08	-0.01	0.04	-0.03				
TLX – Effort	0.10	-0.19	0.18	0.15	0.09	0.07	0.03	-0.04	-0.06	0.22	0.10	0.61	0.48	0.35	0.18			
TLX – Frustration	0.02	-0.10	0.13	0.17	0.25	-0.07	0.06	-0.05	-0.33	0.47	0.18	0.18	0.10	0.38	0.27	0.31		
TLX – Overall	0.13	-0.24	0.20	0.07	0.03	0.06	-0.05	-0.05	-0.05	0.39	0.16	0.68	0.52	0.61	0.41	0.73	0.63	

*Note.* These correlation coefficients were calculated by first averaging the runs in the experiment, and standardizing the data per experiment so that the mean equals zero and the standard deviation equals 1. Correlations of magnitude 0.25 or greater are statistically significant from 0 ( $p < 0.05$ ).  $N = 60$  for the cardiovascular measures due to four missing values in Experiment 2 (Chapter 4).

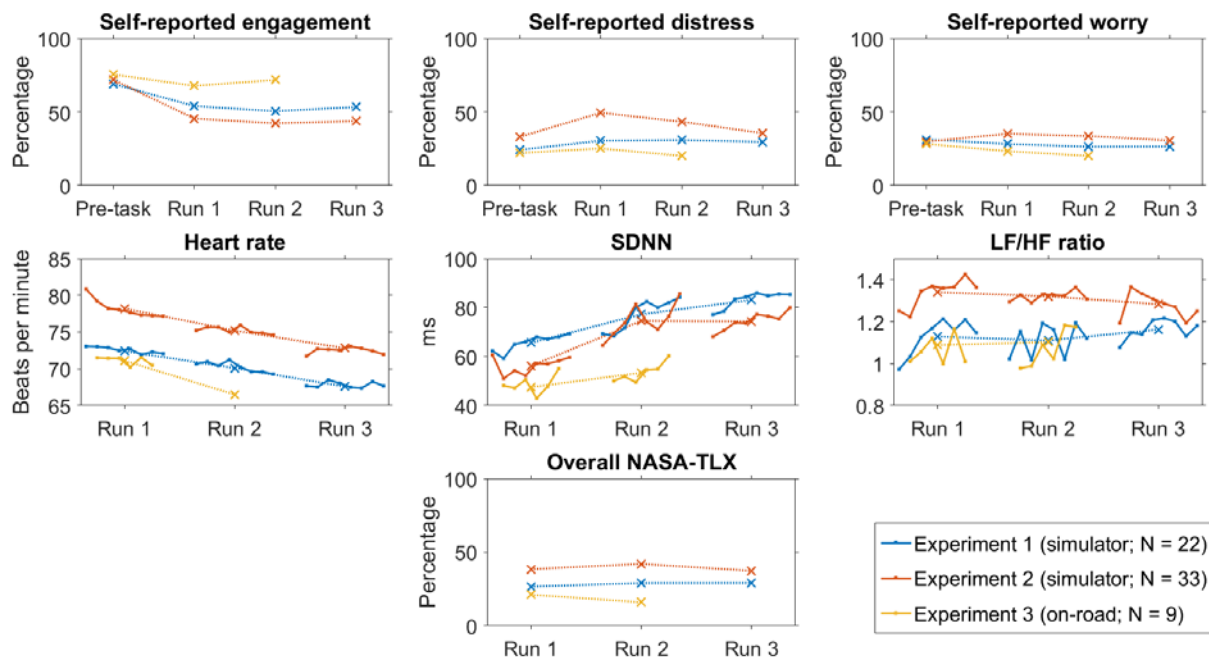
The study described in Chapter 5 measured driver state in an on-road environment. Seven identical measures were compared and showed similar trend lines in heart rate and SDNN, as well as similarly constant self-report scores (Figure 6.2). Nevertheless, the engagement score was higher in the on-road experiment, which may confirm the abovementioned suggestion that drivers in the simulator experiment felt safer due to the absence of harmful consequences in case of a crash. Note, however, that the participants of the on-road experiment described in Chapter 5 were expert drivers, which may explain why they were relatively engaged and non-stressed as compared to the simulator participants. Thus, it remains to be seen how these results translate to regular or novice drivers. Furthermore, the higher results of the Overall NASA-TLX in the second experiment (Chapter 4) are because mental demands were the independent variable in this experiment.

In conclusion, based on the comparisons made in Figure 6.2, it appears that several results from simulator experiments are also representative for an on-road environment. In all three experiments a decline in heart rate, rise in SDNN, and relatively constant worry and overall workload scores were found. The differences between experiments can be attributed to experiment-specific manipulations (e.g., the induced mental demands in Experiment 2), or to the difference in participant pools (e.g., the professional drivers in Experiment 3).

In the simulator experiments in this thesis, a full-sized car (Jaguar XJ saloon) on a fixed-base platform with a relatively wide field-of-view (Figure 6.3) was used, which can be classified as ‘medium fidelity’ (Fisher, Rizzo, Caird, & Lee, 2011). Lower fidelity (e.g., desktop simulators) or higher fidelity (e.g., moving-base simulators) could potentially generate different results, although higher fidelity does not always imply more valid data. Park, Allen, Rosenthal, and Fiorentino (2005) showed that several driver performance measures, such as the number of accidents and speed violations, were similar in simulators with different levels of fidelity. Relatedly, Neubauer, Matthews, Saxby, and Langheim (2010)



argued that stress and fatigue responses are similar in on-road studies as in simulator studies. Whether a simulator with a specific level of fidelity is appropriate depends on the type of research and the sort of research question that is intended to be addressed (see Parkes, 2012 for an overview). Relatedly, one should not simply resort to the highest fidelity simulator available, as the higher the fidelity, the lower the experimental control (Lee, 2004; see Figure 6.4 for a representation).



*Figure 6.2.* The comparisons made in Chapter 5. Top: The Dundee Stress State Questionnaire (DSSQ). Middle: Cardiovascular measures. Bottom: Self-reported overall workload. Experiment 1: Heikoop, De Winter, Van Arem, and Stanton (2017) ( $N = 22$ ; 3 x 40 minutes of simulated driving). Experiment 2: Heikoop, De Winter, Van Arem, and Stanton (submitted-b) ( $N = 33$  [29 for heart rate measures]; 3 x 40 minutes of simulated driving). Experiment 3: Heikoop, De Winter, Van Arem, and Stanton (submitted-a) ( $N = 9$ ; 2 x 32 minutes of on-road driving). Note: The DSSQ used in Experiment 2 is the short version of the DSSQ (see Matthews, Emo, & Funke, 2005).



Figure 6.3. The Southampton University Driving Simulator (SUDS) during the experiment.

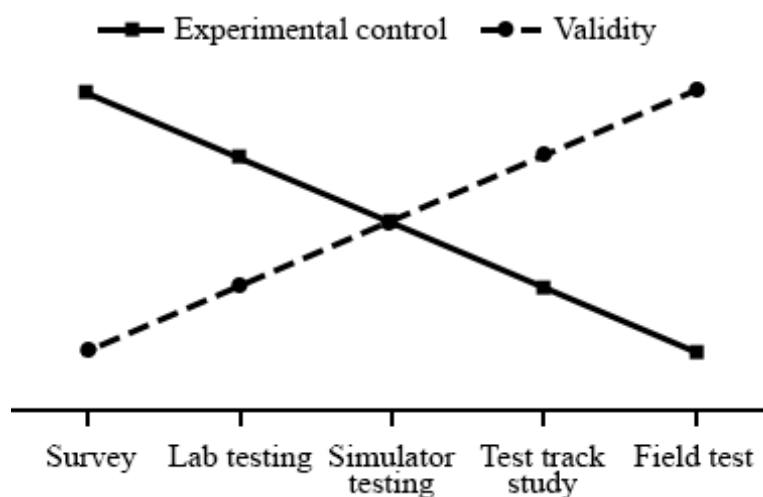


Figure 6.4. Relationship between validity and experimental control of a driving simulator. Copyright: Michael Manser. Derived from: <https://www.fhwa.dot.gov/advancedresearch/pubs/14077/001.cfm>

## Validity of Psychophysiological Measures

The results from the experiments performed in this thesis show that the psychophysiological measures were more sensitive to time-on-task effects than to the type-of-task effects. Although effects were found between conditions (i.e., significantly lower LF/HF ratio in the voluntary condition [Chapter 3], and higher heart rate and LF/HF ratio, and decreased SDNN with increased mental demands [Chapter 4]), the effects between runs were

considerably more distinct, with heart rate declining between runs 1 and 2 (Chapter 5) and also 3 (Chapter 3 & 4) and SDNN (Chapter 3, 4, & 5) rising over time. Moreover, the decline in heart rate could even be seen *between* runs (Figure 6.2), which may suggest that heart rate does not reflect the effects of driving in a platoon for prolonged periods per se, but merely the effects of sitting still for prolonged periods.

Furthermore, from Table 5.4 it could be seen that large individual differences exist regarding the participants' physiological responses, making it difficult to investigate reliable trends in physiological responses. This can be attributed to differences in personal traits, which is in line with Mulder (1992), who found that different people may invest different levels of effort into a task, and as a result, a physiological measure of a person may not represent the level of task complexity, but rather the person's effort in trying to complete the task. Thus, psychophysiological measures such as heart rate variability may be more susceptible to human trait, such as emotional regulation and general health, rather than to human state (see Berntson et al., 1997; McCraty & Shaffer, 2015; Thayer, Åhs, Fredrikson, Sollers, & Wager, 2012 for reviews). This, together with the fact that heart rate variability has a substantial amount of confounding factors, such as blood pressure, physical workload, respiratory rate, and circadian rhythms (e.g., Bernardi, Valle, Coco, Calciati, & Sleight, 1996; Huikuri et al., 1994; Mulder, 1992), implies that although the measurement of heart rate measures can be relatively simple, interpretation and analysis of the results is difficult. Especially in the relatively dynamic environment of a car, which was the case in the experiments performed in this thesis, this could have played a role.

## **Future Design and Implementations**

This thesis has shed light on the psychological effects of the role of a driver in an automated platoon when he or she is driving within that platoon for extended periods of time.

The results from the experiments performed in this thesis show that although drivers appear to be able to remain relatively vigilant (Chapters 3 & 5), the increase in fatigue (as measured by a decrease in heart rate over time; Chapters 3–5), the disengagement from the driving task relative to the pre-task baseline measurement (Chapters 3–5), the apparent digression towards increasingly complex non-self-correcting mental models (Chapter 4), and the fact that the drivers were not performing *perfectly* on a simple detection task (Chapter 3 & 5) suggest that a supervising role of a driver within a platoon is currently misplaced. Moreover, the fact that drivers' situation awareness becomes significantly affected by (non-driving) mental task demands introduces the concern of getting back in the loop. Thus, it appears to be best to avoid putting drivers of a platoon in the position where they have to constantly remain vigilant, or be able to quickly regain situation awareness in a monotonous environment, as even though they know they have full responsibility, they tend to "... do really stupid things when they're behind the wheel" (Teller, 2015). This implies that only fully automated platooning appears to be the only real solution that counters all of the issues above. But until full automation is technologically feasible, keeping the driver engaged using interactive human-machine interfaces (HMI) appears to be the best solution.

As mentioned above, one of the main issues within automated platooning is the fact that drivers become disengaged from the driving task. The results from this thesis regarding task disengagement are in coherence with previous research that found that drivers in a platoon reported a wish for a fatigue-countering system during long drives (Cha & Park, 2006). Levitan, Golembiewski, and Bloomfield (1998) suggested drivers in an automated platoon to perform a 'readiness test' prior to control handover. This is to ensure no carry-over effects occur in manual driving (see also Brandenburg & Skottke, 2014; Skottke, Debus, Wang, & Huestegge, 2014). Henceforth, it is suggested that time-on-task should be given at least as much consideration as type-of-task when designing for HMIs that aim to keep the

driver engaged and comfortable, such as a quiz-like type of interface (cf. Bueno et al., 2016; Verwey & Zaidel, 1999). The disengagement from the driving task found in Chapters 3 and 4 (see also Fig. 3) could be attributed to the fact that drivers tend to over-rely on the automation to drive safely (see also Young & Stanton, 2007). This tendency could have been encouraged by the fact that the forward field of view was blocked by the car ahead. Indeed, research in manual driving found that drivers tend to drive closer to trucks than to regular cars (e.g., Brackstone, Waterson, & McDonald, 2009), which is thought to be done to minimize the number of stimuli to attend to ("Ignorance is bliss"; Sayer, Mefford, & Huang, 2000), although there appears to be no consensus on either the results or a rationale for this behaviour (see e.g., Yeung & Wong, 2014 for a short discussion). In summary, the disengagement from the driving task found in this thesis may, in fact, be due to drivers wanting to become disengaged, and overly rely on the automation to drive safely. A solution to overcome this type of disengagement—and consequently increase safety—could, for example, be the implementation of a 'see-through' screen at the back of the vehicle ahead, a technology recently used in truck platoons (Zhang, Wildschut, Willemsen, Alkim, & Martens, 2018).

Further to a disengagement from the driving task, drivers could become underloaded in an automated vehicle, and thereby resort to distracting non-driving tasks, which in turn could lead to a loss of situation awareness (Rogers et al., 2011; Young & Stanton, 2002; see also Chapter 4). An HMI that, for example, implements colour-coded representations of the automation's status and possible threats within the vehicle's vicinity (i.e., green is good, red is bad, when a sensor is defective, road markings are missing, or during adverse weather; see Scholtz, Antonishek, and Young (2004) for a similar approach with a supervisory interface) could be valuable for regenerating drivers' situation awareness (see also Stanton, Dunoyer, & Leatherland, 2011, who proposed a radar display, variations of which now commonly occur in

consumer-market vehicles). Otherwise, the call for full automation (i.e., the complete removal of responsibility of the driver during the platooning drive) is repeated (see e.g., Farber, 1999; Van Nunen, Ploeg, Morales Medina, & Nijmeijer, 2013).

Regarding the continuous build-up of passive fatigue, implementations of automated driving assistance systems (ADAS), such as a ‘drowsiness warning system’ as mentioned by Cha and Park (2006), or a ‘driver drowsiness detector’ (Hancock & Verwey, 1997) could help a driver of a platoon to become aware of their impending dangerous levels of fatigue. Several detection measures, such as PERCLOS (Wierwille, Ellsworth, Wreggit, Fairbanks, & Kirn, 1994), a dead-man switch, or eye- and performance-based detection measures have been suggested to prevent drivers from falling asleep (May & Baldwin, 2009; Wang & Xu, 2016). However, further to countering fatigue, regeneration of (full) situation awareness is required before regaining control of the vehicle. Research has shown that although participants are relatively quick in assessing basic elements of the situation (i.e., level 1 situation awareness), they need over 20 seconds to gain higher levels of awareness, which entails for example relative speeds and geometrics (Lu, Coster, & De Winter, 2017; see Figure 6.5 for an illustration). Indeed, the detection tasks in Chapters 3 and 5 yielded high hit rates. This may have been because the stimuli were high in saliency (much like visual warning signals, e.g., take-over requests), thus requiring low levels of awareness. For a take-over to appropriately succeed, higher levels of situation awareness may be necessary.

A driver-state monitor that is able to determine appropriate regeneration of situation awareness has been achieved is thus also warranted. Therefore, unless a combination of such measures and interfaces is implemented that encompasses *all* facets (i.e., sufficiently countered fatigue to perceive [level 1 SA], appropriate levels of mental workload to comprehend [level 2 SA], and the right amount of attention to the important stimuli to project

[level 3 SA]) necessary for full situation awareness of the driver, human-supervised platooning appears problematic.

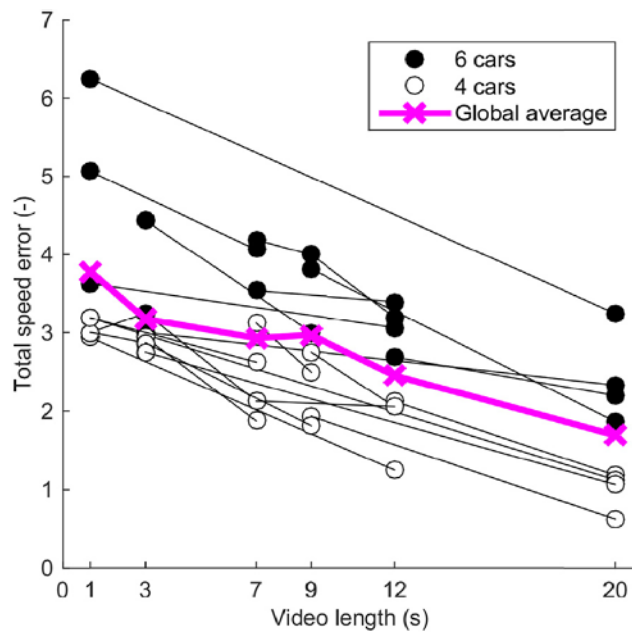


Figure 6.5. Results of a traffic situation recall task (copied from Lu et al., 2017). In the recall task, participants viewed animated video clips showing traffic situations with 4 or 6 surrounding cars. In this particular task, participants were required to watch 1-20 seconds videos of traffic situations, and afterwards indicate whether a car in their surroundings drove slower, faster or at equal speed as the ego vehicle.

In light of the progressive evolution of drivers' mental models seen in the automated platooning scenario (Chapter 4), to avoid *excessive* elaboration on a driver's mental model while driving in an automated platoon, the suggestion is to implement appropriate information about driving with automated driving systems in driver training courses. Providing drivers with a 'ground truth' about the automated system might refrain them from 'fantasizing' about what the automated system can do, and may even result in a diminishing mental model based on what is *not* encountered (Beggiato & Krems, 2013; Beggiato, Pereira, Petzoldt, & Krems, 2015). Since our results did not show that unexperienced situations disappear from a driver's mental model, nor that they digress towards an 'ideal state', in circumstances wherein drivers are unaware of the technology in and capabilities of the automated system (cf. Beggiato & Krems, 2013), it appears drivers could benefit from being informed about the workings of the automated system, in order to get appropriate understanding of its limitations. Moreover, because it is known that people usually do not read their (car) manuals (Mehlenbacher, Wogalter, & Laughery, 2002), it seems vital for drivers to be taught in a mandatory fashion,

and henceforth be updated about recent advances regarding automated driving systems in consumer market vehicles.

As a penultimate point of discussion, in the experiments performed in this thesis we used psychophysiological measures, which proved to be challenging to interpret. It is suggested that future research should go more in depth, assessing only one or two links within the model in a controlled manner, rather than an entire section covering several constructs and links, as was done in this thesis. To be able to pinpoint a single psychological construct, one suggestion is the use of brain imaging techniques, such as electroencephalography (EEG) or functional magnetic resonance imaging (fMRI). For instance, several distinct areas of the brain can be seen to become activated at different stages during a Wisconsin Card Sorting Test, distinguishing, for example, changes in mental workload and attention (Monchi, Petrides, Petre, Worsley, & Dagher, 2001). However, these techniques are no panacea and are likely to be cumbersome when used in a highly dynamic, relatively uncontrolled environment such as a car, as opposed to a laboratory.

The final suggestion that is made in this discussion is to investigate driver's take-over quality, after having driven in an automated platoon for extended periods of time. This thesis assessed the psychological effects of driving *within* an automated platoon but did not investigate manual driving *outside* of an automated platoon. In other words, although this thesis has shown several psychological effects to be significant, it remains to be seen how adverse these effects are to manually controlling the vehicle afterwards. Previous research has shown several carry-over effects after having driven in a platoon for approximately 20 minutes, such as decreased time headway (Brandenburg & Skottke, 2014; Skottke, Debus, Wang, & Huestegge, 2014). Future research could investigate the development of a driver's take-over quality in different scenarios or manoeuvres (e.g., exiting the platoon or intervening in a critical situation, or whilst driving in mild versus heavy traffic), after various time



intervals (e.g., 20 minutes versus 40 minutes versus 80 minutes or longer), and with the addition of a variety of HMIs (e.g., a see-through screen, or a driver-state monitor).

## References

- Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, 10, 229-240. doi: 10.1037/1089-2680.10.3.229
- Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour*, 18, 47-57. doi: 10.1016/j.trf.2012.12.006
- Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F*, 35, 75-84. doi: 10.1016/j.trf.2015.10.005
- Bellet, T., Bailly-Asuni, B., Mayenobe, P., & Banet, A. (2009). A theoretical and methodological framework for studying and modelling drivers' mental representations. *Safety Science*, 47, 1205-1221. doi: 10.1016/j.ssci.2009.03.014
- Bernardi, L., Valle, F., Coco, M., Calciati, A., & Sleight, P. (1996). Physical activity influences heart rate variability and very-low-frequency components in Holter electrocardiograms. *Cardiovascular Research*, 32, 234-237. doi: 10.1016/0008-6363(96)00081-8
- Berntson, G. G., Bigger jr., J. T., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., . . . Van der Molen, M. W. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, 34, 623-648. doi: 10.1111/j.1469-8986.1997.tb02140.x
- Brackstone, M., Waterson, B., & McDonald, M. (2009). Determinants of following headway in congested traffic. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12, 131-142. doi: 10.1016/j.trf.2008.09.003
- Brandenburg, S., & Skottke, E-M. (2014). Switching from manual to automated driving and reverse: Are drivers behaving more risky after highly automated driving? *Proceedings of the IEEE International Conference on Intelligent Transportation Systems (ITSC)* (pp. 2978-2983). Qindao, China.
- Brookhuis, K. A., & De Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis and Prevention*, 42, 898-903. doi: 10.1016/j.aap.2009.06.001
- Bueno, M., Dogan, E., Hadj Selem, F., Monacelli, E., Boverie, S., & Guillaume, A. (2016). How different mental workload levels affect the take-over control after automated driving. *IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. doi: 10.1109/ITSC.2016.7795886
- Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45, 16-28. doi: 10.1037/0003-066X.45.1.16
- Cha, D. W., & Park, P. (2006). Simulator-based Human Factors evaluation of automated highway system. *International Journal of Automotive Technology*, 7, 625-635.
- De Winter, J. C. F., Van Leeuwen, P. M., & Happee, R. (2012). Advantages and disadvantages of driving simulators: A discussion. In A. J. Spink, F. Grieco, O. E.

- Krips, L. W. S. Loijens, L. P. J. J. Noldus & P. H. Zimmerman (Eds.), *Proceedings of Measuring Behavior* (pp. 47-50). Utrecht, The Netherlands, August 28-31, 2012.
- Desmond, P. A., & Matthews, G. S. (2009). Individual differences in stress and fatigue in two field studies of driving. *Transportation Research Part F*, 12, 265-276. doi: 10.1016/j.trf.2008.12.006
- Farber, E. I. (1999). Comments on "Driving behaviour in an emergency situation in the Automated Highway System". *Transportation Human Factors*, 1, 83-85. doi: 10.1207/sthf0101\_8
- Fisher, D. L., Rizzo, M., Caird, J., & Lee, J. D. (2011). *Handbook of driving simulation for engineering, medicine, and psychology*. Boca Raton, FL: CRC Press, Taylor & Francis Group.
- Hallvig, D., Anund, A., Fors, C., Kecklund, G., Karlsson, J. G., Wahde, M., & Åkerstedt, T. (2013). Sleepy driving on the real road and in the simulator - A comparison. *Accident Analysis and Prevention*, 50, 44-50. doi: 10.1016/j.aap.2012.09.033
- Hancock, P. A., & Verwey, W. B. (1997). Fatigue, workload and adaptive driver systems. *Accident Analysis and Prevention*, 29, 495-506. doi: 10.1016/S0001-4575(97)00029-8
- Healey, J. A., & Picard, R. W. (2004). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6, 156-166. doi: 10.1109/TITS.2005.848368
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2017). Effects of platooning on signal-detection performance: A driving simulator study. *Applied Ergonomics*, 60, 116-127. doi: 10.1016/j.apergo.2016.10.016
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (submitted-a). Acclimatizing to automation: driver workload and stress during partially automated car following in real traffic.
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (submitted-b). Effects of mental demands on situation awareness during automated platooning: A driving simulator study.
- Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, 45, 73-93. doi: 10.1016/S0301-0511(96)05223-4
- Huikuri, H. V., Niemela, M. J., Ojala, S., Rantala, A., Ikaheimo, M. J., & Airaksinen, K. E. J. (1994). Circadian rhythms of frequency domain measures of heart rate variability in healthy subjects and patients with coronary artery disease effects of arousal and upright posture. *Circulation*, 90, 121-126. doi: 10.1161/01.CIR.90.1.121
- Lal, S. K. L., & Craig, A. (2000). Driver fatigue: Psychophysiological effects 4th *International Conference on Fatigue and Transportation* Fremantle, Western Australia.
- Le, H., Schmidt, F. L., Harter, J. K., & Lauver, K. J. (2010). The problem of empirical redundancy of constructs in organizational research: An empirical investigation. *Organizational Behavior and Human Decision Processes*, 112, 112-125. doi: 10.1016/j.obhdp.2010.02.003
- Lee, J. D. (2004). Simulator fidelity: How low can you go? *Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA.
- Leviton, L., Golembiewski, G., & Bloomfield, J. R. (1998). Human Factors issues for Automated Highway Systems. *ITS Journal – Intelligent Transportation Systems Journal: Technology, Planning, and Operations*, 4, 21-47. doi: 10.1080/10248079808903735
- Llaneras, R. E., Salinger, J., & Green, C. A. (2013). Human Factors issues associated with limited ability autonomous driving systems: Drivers' allocation of visual attention to

- the forward road. *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (pp. 92-98). Bolton Landing, NY.
- Lu, Z., Coster, X., & De Winter, J. C. F. (2017). How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving. *Applied Ergonomics*, 60, 293-304. doi: 10.1016/j.apergo.2016.12.003
- Ma, R., & Kaber, D. B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics*, 35, 939-953. doi: 10.1016/j.ergon.2005.04.002
- Matthews, G. S. (2002). Towards a transactional ergonomics for driver stress and fatigue. *Theoretical Issues in Ergonomics Science*, 3, 195-211. doi: 10.1080/14639220210124120
- Matthews, G. S., Emo, A. K., & Funke, G. J. (2005). *A short version of the Dundee Stress State Questionnaire*. Paper presented at the Twelfth Meeting of the International Society for the Study of Individual Differences, Adelaide, Australia.
- Matthews, G. S., Joyner, L., Gilliland, K., Campbell, S., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state 'Big Three'? In I. Mervielde, I. J. Deary, F. De Fruyt & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 335-350). Tilburg: Tilburg University Press.
- May, J. M., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation Research Part F*, 12, 218-224. doi: 10.1016/j.trf.2008.11.005
- McCraty, R., & Shaffer, F. (2015). Heart rate variability: new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. *Global Advances in Health and Medicine*, 4, 46-61. doi: 10.7453/gahmj.2014.073
- Mehlenbacher, B., Wogalter, M. S., & Laughery, K. R. (2002). On the reading of product owners' manuals: Perceptions and product complexity. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, 730-734. doi: 10.1177/154193120204600610
- Mehler, B. (2015). *Recording heart rate & electrodermal activity as measures of cognitive load in the driving environment*.
- Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, 21, 7733-7741.
- Mulder, L. J. M. (1992). Measurement and analysis methods of heart rate and respiration for use in applied environments. *Biological Psychology*, 34, 205-236. doi: 10.1016/0301-0511(92)90016-N
- Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality*, 41, 1213-1233. doi: 10.1016/j.jrp.2007.02.003
- Neubauer, C., Matthews, G. S., Saxby, D., & Langheim, L. (2010). Simulator methodologies for investigating fatigue and stress in the automated vehicle. *Advances in Transportation Studies: An International Journal, Special Issue*, 7-18. doi: 10.4399/97888548416592
- Omae, M., Hashimoto, N., Sugamoto, T., & Shimizu, H. (2005). Measurement of driver's reaction time to failure of steering controller during automatic driving. *Review of Automotive Engineering*, 26, 213-215.
- Park, G. D., Allen, R. W., Rosenthal, T. J., & Fiorentino, D. (2005). Training effectiveness: How does driving simulator fidelity influence driver performance? *Proceedings of the*

- Human Factors and Ergonomics Society Annual Meeting*, 49, 2201-2205. doi: 10.1177/154193120504902518
- Parkes, A. (2012). The essential realism of driving simulators for research and training. In N. Gkikas (Ed.), *Automotive ergonomics: Driver-vehicle interaction* (pp. 133-154). Boca Raton, FL: CRC Press.
- Reed, M. P., & Green, P. A. (1999). Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task. *Ergonomics*, 42, 1015-1037. doi: 10.1080/001401399185117
- Rogers, M., Zhang, Y., Kaber, D., & Liang, Y. (2011). The effects of visual and cognitive distraction on driver situation awareness. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics. EPCE 2011. Lecture Notes in Computer Science* (pp. 6781). Berlin, Heidelberg: Springer.
- Saxby, D. J., Matthews, G. S., Hitchcock, E. M., & Warm, J. S. (2007). Development of active and passive fatigue manipulations using a driving simulator. *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, 1237-1241. doi: 10.1177/145193120705101839
- Sayer, J. R., Mefford, M. L., & Huang, R. (2000). The effect of lead-vehicle size on driver following behavior. (Report No. UMTRI-2000-15), The University of Michigan Transportation Research Institute: Ann Arbor, MI.
- Scholtz, J., Antonishek, B., & Young, J. (2004). Evaluation of a human-robot interface: Development of a situational awareness methodology. *Proceedings of the 37th Annual Hawaii International Conference on Systems Sciences*. doi: 10.1109/HICSS.2004.1265327
- Shaffer, J. A., DeGeest, D., & Li, A. (2017). Tackling the problem of construct proliferation: A guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods*, 19, 80-110. doi: 10.1177/1094428115598239
- Skottke, E-M., Debus, G., Wang, L., & Huestegge, L. (2014). Carryover effects of highly automated convoy driving on subsequent manual driving performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56, 1272-1283. doi: 10.1177/0018720814524594
- Stanton, N. A., Dunoyer, A., & Leatherland, A. (2011). Detection of new in-path targets by drivers using stop & go adaptive cruise control. *Applied Ergonomics*, 42, 592-601. doi: 10.1016/j.apergo.2010.08.016
- Stanton, N. A., & Marsden, P. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science*, 24, 35-49. doi: 10.1016/S0925-7535(96)00067-7
- Stanton, N. A., & Young, M. S. (2000). A proposed psychological model of driving automation. *Theoretical Issues in Ergonomics Science*, 1, 315-331. doi: 10.1080/14639220052399131
- Teller, A. (2015). *Moonshots and reality*. Keynote Presentation given at South by Southwest (SXSW), Austin, TX, March 17.
- Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews*, 36, 747-756. doi: 10.1016/j.neubiorev.2011.11.009
- Van Nunen, E., Ploeg, J., Morales Medina, A., & Nijmeijer, H. (2013). Fault tolerancy in Cooperative Adaptive Cruise Control. *16th International IEEE Annual Conference on Intelligent Transportation Systems (ITSC 2013)*, 1184-1189. doi: 10.1109/ITSC.2013.6728393

- Verwey, W. B., & Zaidel, D. M. (1999). Preventing drowsiness accidents by an alertness maintenance device. *Accident Analysis and Prevention*, 31, 199-211. doi: 10.1016/S0001-4575(98)00062-1
- Wang, X., & Xu, C. (2016). Driver drowsiness detection based on non-intrusive metrics considering individual specifics. *Accident Analysis & Prevention*, 95, Part B, 350-357. doi: 10.1016/j.aap.2015.09.002
- Wickens, C. D. (2002). Situation awareness and workload in aviation. *Current Directions in Psychological Science*, 11, 128-133. doi: 10.1111/1467-8721.00184
- Wierwille, W. W., Ellsworth, L. A., Wreggit, S. S., Fairbanks, R. J., & Kirn, C. L. (1994). *Research on vehicle-based driver status/performance monitoring: Development, validation, and refinement of algorithms for detection of driver drowsiness*. (Report No. DOT HS 808 247), National Highway Traffic Safety Administration, Washington, DC.
- Yeung, J. S., & Wong, Y. D. (2014). The effect of road tunnel environment on car following. *Accident Analysis and Prevention*, 70, 100-109. doi: 10.1016/j.aap.2014.03.014
- Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors: The Journal of Human Factors and Ergonomics Society*, 44, 365-375. doi: 10.1518/0018720024497709
- Young, M. S., & Stanton, N. A. (2007). Back to the future: Brake reaction times for manual and automated vehicles. *Ergonomics*, 50, 46-58. doi: 10.1080/00140130600980789
- Zhang, B., Wildschut, E. S., Willemsen, D. M. C., Alkim, T., & Martens, M. H. (2018). The effect of see-through truck on driver monitoring patterns and responses to critical events in truck platooning. In N. A. Stanton (Ed.), *Advances in Human Aspects of Transportation*: Springer International Publishing.

## **Acknowledgements**

This section will contain thank words to those that deserve to be thanked. By their selfless effort, unparalleled supervision, and augmenting critique, they elevated me to levels exceeding my own expectation. The first person I would like to thank is dr. Joost de Winter, who supervised me throughout the three years I was doing my PhD under his wing. Without him, I would have certainly not been able to get anywhere near the position where I am at now. The numerous emails we exchanged, progressively improving my written words, enabled me to get work publishable, and in the end, a thesis worth reading through. This leads me to the next person worthy of thanks, prof. Neville Stanton. He believed in me in times where several other companies and corporations did not believe me to be a credible employee. Even when I almost had to go back home empty-handed because I wasn't allowed to do a PhD at the University of Southampton, he made all the effort for me to stay there, and become an external PhD student somewhere else. And this is where the final of three supervisors comes into my words of thanks, prof. Bart van Arem, who made it possible for me to become an external PhD student at the Delft University of Technology. Without his belief in my good continuance, I would have had nothing. So, once more, eternal thanks are due to all three of my supervisors.

Then rests me to thank everyone involved with the HFAuto project. First, the leader of the project, dr. Riender Happee, who made it possible for us to work in a wonderful project, and ensured successful proceeding of the project as a whole. The experiences I gained during this project were priceless, and the fact that I have met so many incredible people will also never leave my mind. It was a pleasure to work with such like-minded people, and the adventures, talks and experiences not directly related to work or academia have also been of great joy to me.

## Acknowledgements

This leads me to the colleagues, and friends, at the University of Southampton. It was amazingly good to have worked with all of them. The atmosphere created there was wonderful, and encouraged me in more ways than one could think of. Certainly a great amount of the reasons I made my PhD a success is due to their accompaniment. How many people were there to help me in times where and on topics with which I struggled? Definitely too many to count. Some people helped me with hard- or software issues. Others helped me think of research methodologies. And again others helped me review my work, and even more others were there to help me power through the endless woods of bureaucracy.

Also credit is due to the people from Jaguar/LandRover, who enabled me to perform an on-road study at their location. This is of course in our domain a priceless experience, so many thanks to those who made that happen.

Shall I almost forget my amazing, lovely, beautiful wife, Siswa? Definitely not! The amount of sorrows she had to deal with is inestimable. And the amount of support I received from her is unbelievable. She is definitely a big part in my success. In fact, at several occasions, she completely ignored her own problems, and joined me in tackling mine. Maybe, just maybe, or maybe not so maybe, I wouldn't have made it without her.

As a final word of thanks, I would like to acknowledge my parents, family and friends, and all other people related to me who were often keen to know what I did, where I was at, and how I did. In many ways they were not able to understand, as the world of academia is clearly something different, but either way, their interest has also been a thriving force for me.

## **Curriculum Vitae**

### **Personalia**

Daniël D. Heikoop  
31 Waldegrave Close  
Southampton SO199RY  
United Kingdom

Tel: +44 7483804542  
E-mail: [ddheikoop@gmail.com](mailto:ddheikoop@gmail.com)

Marital status: Married  
Nationality: Dutch

Place of birth: Leiderdorp  
Date of birth: 15 June 1986



### **Education**

2014-current **TU Delft, Delft**  
PhD: Human Behaviour During Highly Automated Driving  
2011-2012 **University Utrecht, Utrecht (diploma)**  
Master Applied Cognitive Psychology  
2006-2011 **University Utrecht, Utrecht (diploma)**  
Bachelor Cognitive Artificial Intelligence  
2005-2006 **TU Delft, Delft**  
Architecture (switched)  
1998-2005 **Groene Hart Lyceum, Alphen aan den Rijn**  
Pre-university education

### **Job experience**

2014-current **University of Southampton**  
Senior Research Assistant  
2013-2014 **Infostrada Sports Group at Nieuwegein**  
Volunteer at the Biography Research Project  
2012 **Netherlands Institute for the Study of Crime and Law Enforcement (NSCR) at Amsterdam & Scientific Research- and Documentationcentre at the Hague (WODC)**  
Internship (5.5 months)  
2011-2012 **Intergo BV at Utrecht**  
Internship (2.5 months)  
2003-2014 **Multiple side runs**



## Languages

---

Dutch	Native
English	Fluent
German	Good
French	Basic

## Personality

---

Self-sufficient, team player, works well under pressure, result-driven, accurate, honest, friendly, flexible, ambitious, competitive

## Theses

---

Driver psychology during automated platooning	PhD thesis
Characterizing male adult convicts through psychophysiological measures	Master thesis
Processing emotional stimuli in the visual system using the low road	Bachelor thesis
Physiological measures and measuring systems for objectively assess mental workload (Original title: <i>Fysiologische maten en meetsystemen voor het objectief bepalen van mentale werkbelasting</i> )	Internship report

## Publications

---

### Journal publications

- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2017). Effects of platooning on signal-detection performance, workload, stress and fatigue: A driving simulator study. *Applied Ergonomics*, 60, 116-127. doi:10.1016/j.apergo.2016.10.016
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2016). Psychological construct in driving automation: A consensus model and critical comment on construct proliferation. *Theoretical Issues in Ergonomics Science*, 17(3), 284-303. doi:10.1080/1463922X.2015.1101507

### Conference presentations

- Heikoop, D. D., Lu, Z., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2017). Temporal effects in automated driving: the decline and rise of attentional state. *Abstract accepted to Road Safety & Simulation International Conference (RSS2017)*, Delft, The Netherlands.
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2016). Highly automated platooning: Effects on mental workload, stress and fatigue. *Abstract accepted to Ergonomics and Human Factors 2016 (EHF2016)*, Staverton Park in Daventry, Northamptonshire.
- Heikoop, D. D., De Winter, J. C. F., Van Arem, B., & Stanton, N. A. (2015). Psychological construct in driving automation: A consensus model and critical comment on construct proliferation. *Presentation at the 6<sup>th</sup> International Conference on Applied Human Factors and Ergonomics*, Las Vegas, NV.

## **The 10 Propositions about my PhD Thesis**

These propositions are regarded as opposable and defensible, and have been approved as such by the (co)promotor(s) dr. ir. Joost C. F. de Winter, prof. dr. ir. Bart van Arem, and prof. dr. Neville A. Stanton

1. Automated platooning should only be pursued with full automation.
2. *If* human-supervised automated driving is used, meaningful human control will be needed to overcome Human Factors issues.
3. *If* human-supervised automated driving is used, ‘RTFM’ must have a dominant role in driver training.
4. In automated driving, we must design for the 99th percentile, not for the mean, of psychological/behavioural responses.
5. Eye tracking for driver-state assessment is viable, but still needs massive improvement.
6. The current psychological literature suffers from construct proliferation.
7. In the domain of ‘Human Factors of automated driving’, individual differences must be the core topic to be studied.
8. In the past, Human Factors research in aviation automation was much ahead of Human Factors research in automated driving. This situation is now reversed.
9. Tesla’s Ludicrous mode is of greater safety concern than its Autopilot mode.
10. The ‘time on task’ has stronger physiological effects than the ‘type of task’ that drivers are performing in an automated car.

## Appendix A: Experimental materials of Chapter 2

### Appendix A.1: Reference list for Table 3.

- [1] Bach, K. M., Jæger, M. G., Skov, M. B., & Thomassen, N. G. (2009). Interacting with in-vehicle systems: Understanding, measuring, and evaluating attention. *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, 453-462. doi:10.1145/1671011.1671070
- [2] Beede, K. E., & Kass, S. J. (2006). Engrossed in conversation: The impact of cell phones on simulated driving performance. *Accident Analysis & Prevention*, 38, 415-421. doi:10.1016/j.aap.2005.10.015
- [3] Bellet, T., Bailly-Asuni, B., Mayenobe, P., & Banet, A. (2009). A theoretical and methodological framework for studying and modelling drivers' mental representations. *Safety Science*, 47, 1205-1221. doi:10.1016/j.ssci.2009.03.014
- [4] Bishop, R. (2005). Intelligent vehicle R&D: A review and contrast of programs worldwide and emerging trends. *Annales Des Télécommunications*, 60, 228-263. doi:10.1007/BF03219820
- [5] Brookhuis, K. A., De Waard, D., & Jansen, W. H. (2001). Behavioural impacts of advanced driver assistance systems-An overview. *European Journal of Transport and Infrastructure Research*, 1, 245-253.
- [6] Carsten, O. M. J., & Nilsson, L. (2001). Safety assessment of driver assistance systems. *European Journal of Transport and Infrastructure Research*, 1, 225-243.
- [7] Collet, C., Petit, C., Priez, A., & Dittmar, A. (2005). Stroop color-word test, arousal, electrodermal activity and performance in a critical driving situation. *Biological Psychology*, 69, 195-203. doi:10.1016/j.biopsycho.2004.07.003
- [8] Collet, C., Clarion, A., Morel, M., Chapon, A., & Petit, C. (2009). Physiological and behavioural changes associated to the management of secondary tasks while driving. *Applied Ergonomics*, 40, 1041-1046. doi:10.1016/j.apergo.2009.01.007
- [9] Collet, C., Guillot, A., & Petit, C. (2010). Phoning while driving II: A review of driving conditions influence. *Ergonomics*, 53, 602-616. doi:10.1080/00140131003769092
- [10] Coughlin, J. F., Reimer, B., & Mehler, B. (2009). *Driver wellness, safety & the development of an awarecar*. Cambridge, MA: MIT AgeLab.
- [11] Desmond, P. A., Hancock, P. A., & Monette, J. L. (1998). Fatigue and automation-induced impairments in simulated driving performance. *Transportation Research Record*, 1628, 8-14. doi:10.3141/1628-02
- [12] Desmond, P. A., & Matthews, G. S. (2009). Individual differences in stress and fatigue in two field studies of driving. *Transportation Research Part F*, 12, 265-276. doi:10.1016/j.trf.2008.12.006
- [13] Van Driel, C. (2007). *Driver support in congestion: An assessment of user needs and impacts on driver and traffic flow*. (Doctoral dissertation), Universiteit Twente.
- [14] Flemisch, F. O., Kelsch, J., Löper, C., Schieben, A., Schindler, J., & Heesen, M. (2008). *Cooperative control and active interfaces for vehicle assistance and automation*. Paper presented at the FISITA World Automotive Congress, Munich, Germany.
- [15] Funke, G., Matthews, G. S., Warm, J. S., & Emo, A. K. (2007). Vehicle automation: A remedy for driver stress? *Ergonomics*, 50, 1302-1323. doi:10.1080/00140130701318830
- [16] Hoc, J.-M., Mars, F., Milleville-Pennel, I., Jolly, É., Netto, M., & Blosseville, J.-M. (2006). Human-machine cooperation in car driving for lateral safety: Delegation and mutual control. *Le Travail Humain*, 69, 153-182. doi:10.3917/th.692.0153

- [17] Hoc, J.-M., Young, M. S., & Blosseville, J.-M. (2009). Cooperation between drivers and automation: Implications for safety. *Theoretical Issues in Ergonomics Science*, 10, 135-160. doi:10.1080/14639220802368856
- [18] 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. *Transportation Research Part C*, 30, 116-125. doi:10.1016/j.trc.2013.02.008
- [19] Kazi, T. A., Stanton, N. A., Walker, G. H., & Young, M. S. (2007). Designer driving: Drivers' conceptual models and level of trust in adaptive cruise control. *International Journal of Vehicle Design*, 45, 339-360. doi:10.1504/IJVD.2007.014909
- [20] Kim, Y.-S. (2001). *Effects of driver, vehicle, and environment characteristics on collision warning systems design*. (Master's thesis), Linköping Institute of Technology.
- [21] Landau, K. (2002). Usability criteria for intelligent driver assistance systems. *Theoretical Issues in Ergonomics Science*, 3, 330-345. doi:10.1080/14639220210137748
- [22] Ma, R., & Kaber, D. B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics*, 35, 939-953. doi:10.1016/j.ergon.2005.04.002
- [23] Matthews, G. S., Campbell, S. E., Desmond, P. A., Huggins, J., Falconer, S., & Joyner, L. A. (1999). Assessment of task-induced state change: Stress, fatigue and workload components. In M. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- [24] Matthews, G. S. (2002). Towards a transactional ergonomics for driver stress and fatigue. *Theoretical Issues in Ergonomics Science*, 3, 195-211. doi:10.1080/14639220210124120
- [25] Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. M. J. (2012). Highly automated driving, secondary task performance and driver state. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 54, 762-771. doi:10.1177/0018720812442087
- [26] Navarro, J., Mars, F., & Young, M. S. (2011). Lateral control assistance in car driving: Classification, review and future prospects. *IET Intelligent Transport Systems*, 5, 207-220. doi:10.1049/iet-its.2010.0087
- [27] Piao, J., & McDonald, M. (2008). Advanced driver assistance systems from autonomous to cooperative approach. *Transport Reviews: A Transnational Transdisciplinary Journal*, 28, 659-684. doi:10.1080/01441640801987825
- [28] Rauch, N., Kaussner, A., Krüger, H.-P., Boverie, S., & Flemisch, F. O. (2009). The importance of driver state assessment within highly automated vehicles. *Proceedings of the ITS – 16th World Congress and Exhibition on Intelligent Transport Systems and Services*.
- [29] Rudin-Brown, C. M. (2010). 'Intelligent' in-vehicle intelligent transport systems: Limiting behavioural adaptation through adaptive design. *IET Intelligent Transport Systems*, 4, 252-261. doi:10.1049/iet-its.2009.0151
- [30] Schieben, A., Damböck, D., Kelsch, J., Rausch, H., & Flemisch, F. O. (2008). Haptisches Feedback im Spektrum von Fahrerassistenz und Automation. *Proceedings of the 3rd Conference on Active Safety through Driver Assistance*.
- [31] Stanley, L. M. (2006). *Haptic and auditory interfaces as a collision avoidance technique during roadway departures and driver perception of these modalities*. (Doctoral dissertation), Montana State University-Bozeman.

- [32] Stanton, N. A., & Young, M. S. (1998). Vehicle automation and driving performance. *Ergonomics*, 41, 1014-1028. doi:10.1080/001401398186568
- [33] Stanton, N. A., & Young, M. S. (2000). A proposed psychological model of driving automation. *Theoretical Issues in Ergonomics Science*, 1, 315-331. doi:10.1080/14639220052399131
- [34] Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics*, 48, 1294-1313. doi:10.1080/00140130500252990
- [35] Stanton, N. A., Dunoyer, A., & Leatherland, A. (2011). Detection of new in-path targets by drivers using stop & go adaptive cruise control. *Applied Ergonomics*, 42, 592-601. doi:10.1016/j.apergo.2010.08.016
- [36] Stanton, N. A., Salmon, P. M., Walker, G. H., & Jenkins, D. P. (2010). Is situation awareness all in the mind? *Theoretical Issues in Ergonomics Science*, 11, 29-40. doi:10.1080/14639220903009938
- [37] Stanton, N. A., Walker, G. H., Young, M. S., Kazi, T. A., & Salmon, P. M. (2007). Changing drivers' minds: The evaluation of an advanced driver coaching system. *Ergonomics*, 50, 1209-1234. doi:10.1080/00140130701322592
- [38] Vahidi, A., & Eskandarian, A. (2003). Research advances in intelligent collision avoidance and adaptive cruise control. *IEEE Transactions on Intelligent Transportation Systems*, 4, 143-153. doi:10.1109/TITS.2003.821292
- [39] Young, M. S., & Stanton, N. A. (1997). Automotive automation: Investigating the impact on drivers' mental workload. *International Journal of Cognitive Ergonomics*, 1, 325-336.
- [40] Young, M. S., Stanton, N. A., & Harris, D. (2007). Driving automation: Learning from aviation about design philosophies. *International Journal of Vehicle Design*, 45, 323-338. doi:10.1504/IJVD.2007.014908
- [41] Warm, J. S., Matthews, G. S., & Finomore, V. S. (2008). Vigilance, workload and stress. In P. A. Hancock & J. Szalma (Eds.), *Performance under Stress* (pp. 225-241). Aldershot: Ashgate Publishing Limited.
- [42] Engström, J., & Hollnagel, E. (2007). A general conceptual framework for modelling behavioural effects of driver support functions. In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments: Critical issues in driver interactions with intelligent transport systems* (pp. 61-84). London: Springer.
- [43] Saxby, D. J., Matthews, G. S., Hitchcock, E. M., & Warm, J. S. (2007). Development of active and passive fatigue manipulations using a driving simulator. *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, 1237-1241. doi:10.1177/145193120705101839

## **Appendix B: Experimental Materials of Chapter 3**

### **Appendix B.1: Experiment instruction sheet, Run 1, Condition No Task.**

#### **Experiment Instructions**

First of all, thank you for taking part in this experiment. Please read the instructions carefully.

#### **Before the experiment**

You are now seated in our Southampton University Driving Simulator (SUDS). This simulator is able to simulate manual, as well as automated driving amongst a variety of scenarios.

Now you will be asked to read and fill out a few questionnaires. These are:

- Participant Information Sheet
  - o This will explain what the research is about and what can and will happen before, during and after the experiment
- Consent Form
  - o This is to ensure you understand what your participation entails
- ECG Electrode placement
  - o This is to inform you of how and where the ECG electrodes will be placed
- General Information
  - o This is to get some general information, solely for research purposes
- State Questionnaire
  - o This is to assess your current state, solely for research purposes

Furthermore, the researcher will monitor your eye- and head movements by means of a remote eye tracker. Some calibrations will have to be made, and therefore your cooperation is appreciated. Also some of your biometric measures, such as your heart rate, will be measured. For this, some electrodes have to be placed on you, as explained in the Information Sheet.

#### **During the experiment**

You will take part in a simulated drive in a highly automated platoon. This will mean you will be part of a group of highly automated vehicles who will drive very close together.

Driving in a highly automated vehicle means that the vehicle automation will take care of both the speed and the steering, so, theoretically, you don't have to do a thing! But an incident is always possible, so you still have to pay attention. The drive will take place along a highway with mild traffic. This first drive (out of three) will be around 40 minutes.

Your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears. Try to avoid any accidents at all times.

**After the experiment**

You will be asked to fill out two questionnaires:

- State Questionnaire
  - This is to assess your current state after your drive
- NASA-TLX
  - This is to assess your experiences about your drive

Then you will receive the next instructions for your second drive.

**Appendix B.2: Experiment instruction sheet, Run 2/3, Condition No Task.**

**Experiment Instructions**

You will now take part in the next simulated drive. This time, your only task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears. Try to avoid any accidents at all times.

After the experiment you will be asked to fill out the same two questionnaires again:

- State Questionnaire
- NASA-TLX



### **Appendix B.3: Experiment instruction sheet, Run 1, Condition Voluntary, Part: During the experiment.**

#### **During the experiment**

You will take part in a simulated drive in a highly automated platoon. This will mean you will be part of a group of highly automated vehicles who will drive very close together.

Driving in a highly automated vehicle means that the vehicle automation will take care of both the speed and the steering, so, theoretically, you don't have to do a thing! But an incident is always possible, so you still have to pay attention. The drive will take place along a highway with mild traffic. This first drive (out of three) will be around 40 minutes.

Your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears.

Furthermore, you are free to undertake any activity during the drive, as long as you try to avoid any accidents at all times. Examples can be reading, drinking, eating, counting red cars, or even sleep, as long as you feel you are able to intervene at all times.

**Appendix B.4: Experiment instruction sheet, Run 2/3, Condition Voluntary.**

**Experiment Instructions**

You will now take part in the next simulated drive. This time, your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears.

Furthermore, you are free to undertake any activity during the drive, as long as you try to avoid any accidents at all times. Examples can be reading, drinking, eating, counting red cars, or even sleep, as long as you feel you are able to intervene at all times.

After the experiment you will be asked to fill out the same two questionnaires again:

- State Questionnaire
- NASA-TLX

### **Appendix B.5: Experiment instruction sheet, Run 1, Condition Detection Task, Part: During the experiment.**

#### **During the experiment**

You will take part in a simulated drive in a highly automated platoon. This will mean you will be part of a group of highly automated vehicles who will drive very close together.

Driving in a highly automated vehicle means that the vehicle automation will take care of both the speed and the steering, so, theoretically, you don't have to do a thing! But an incident is always possible, so you still have to pay attention. The drive will take place along a highway with mild traffic. This first drive (out of three) will be around 40 minutes.

Your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears. Furthermore, you are asked to count the **red** cars during the ride. These can be either on your direction, or the opposite direction. Press the handheld button whenever you see one. Try to avoid any accidents at all times.

**Appendix B.6: Experiment instruction sheet, Run 2/3, Condition Detection Task.**

**Experiment Instructions**

You will now take part in the next simulated drive. This time, your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears. Furthermore, you are asked to count the **red** cars during the ride. These can be either on your direction, or the opposite direction. Press the handheld button whenever you see one. Try to avoid any accidents at all times.

After the experiment you will be asked to fill out the same two questionnaires again:

- State Questionnaire
- NASA-TLX

## **Appendix B.7: Participant information sheet.**

### **Participant Information Sheet**

**Study Title:** Long-Term Effects of Highly Automated Platoons

**Researcher:** Daniel D. Heikoop

**Ethics number:** 13967

**Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.**

#### **What is the research about?**

This study is part of a larger international research project called HF-AUTO. I am interested in the long-term effects of highly automated driving on driving behaviour. This research is funded by the Marie Curie-Skłodowska Actions.

#### **Why have I been chosen?**

You have been approached to take part in this study as part of an effort to recruit people who have a driving license, and are at least in adequate physical and mental health.

#### **What will happen to me if I take part?**

Should you choose to take part, you will be asked to participate in three sessions of driving in the University of Southampton's Driving Simulator. The three sessions will all involve driving a similar route on a highway within a platoon of highly automated vehicles, each lasting approximately 40 minutes, with one short questionnaire to be filled out before, and two short questionnaires to be filled out after each session.

In total, the experiment should last around three hours. The questionnaires will ask about the amount of workload you feel you are under for each driving trial and about your feelings and thoughts at the moment. Furthermore your heart rate will be measured during the ride by electrodes placed on the upper chest- and back region, and your eye movements will be measured by means of a remote eye tracker. Both measurements will not hurt in any sense, and discretion is being considered at all times.

#### **Are there any benefits in my taking part?**

While individual benefit may be limited, your participation will help us to build an understanding of the long-term effects of highly automated driving in platoons on driving behaviour. It is hoped that the results achieved in this research will be used in the development of driving automation in commercially available road vehicles.

#### **Are there any risks involved?**

Driving in a simulator can cause motion sickness in some people, though not all. Some suggest around 10% of people experience symptoms such as nausea and dizziness. Should this occur, participation will be stopped.

Beyond the risk of motion sickness, this study does not carry with it any significant risks additional those you would experience in normal, day-to-day life.

#### **Will my participation be confidential?**

No identifying data will be released to anyone other than the main investigator. Consent forms will be stored securely in a locked cabinet, with the primary investigator having the only key.

#### **What happens if I change my mind?**

You are completely free to end your participation at any time. You are not obliged in any way to continue with the session, or even begin the session, should you so choose.

#### **What happens if something goes wrong?**

In the unlikely case of concern or complaint, please contact the Research Governance Office at the University of Southampton).

#### **Where can I get more information?**

If you would like more information, please feel free to email me, the main investigator, at D.D.Heikoop@soton.ac.uk.

**Appendix B.8: Consent form.**

**CONSENT FORM**

**Study title:** Long-Term Effects of Highly Automated Platoons

**Researcher name:** Daniel D. Heikoop

**Study reference:** -

**Ethics reference:** 13967

*Please initial the box(es) if you agree with the statement(s):*

I have read and understood the information sheet (insert date /version no. of participant information sheet) and have had the opportunity to ask questions about the study.

☐

I agree to take part in this research project and agree for my data to be used for the purpose of this study

☐

I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected

☐

I understand that my participation entails the monitoring of heart rate measures, and I am happy for it to be so

☐

**Data Protection**

*I understand that information collected about me during my participation in this study will be stored on a password protected computer and that this information will only be used for the purpose of this study. All files containing any personal data will be made anonymous.*

Name of participant (print name).....

Signature of participant.....

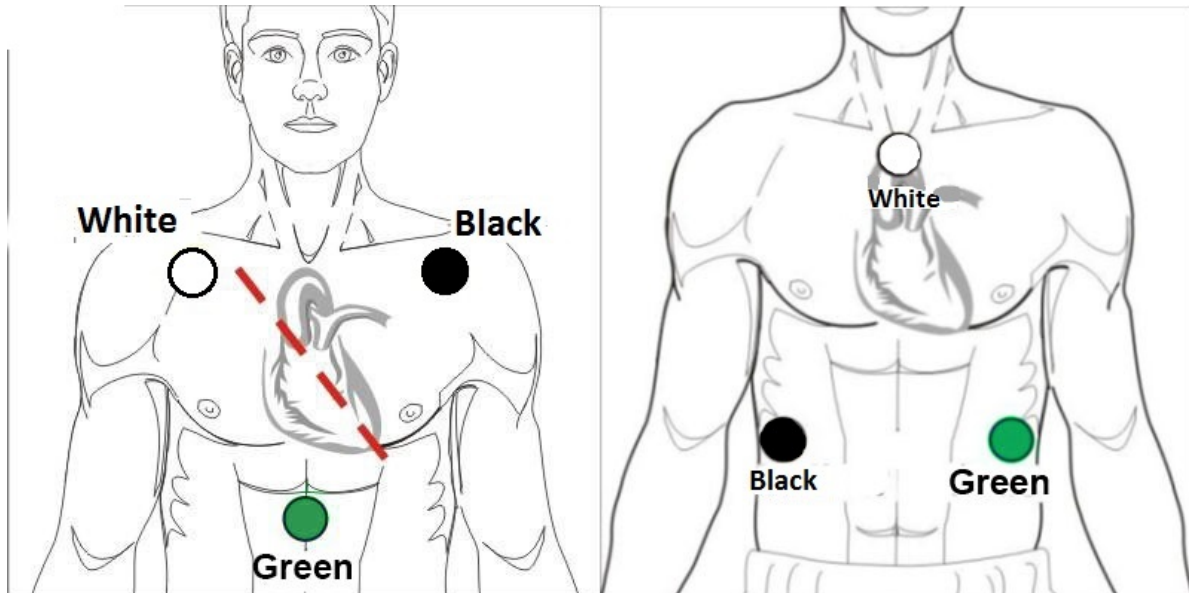
Date.....

**Appendix B.9: ECG electrode placement information sheet.**

**ECG electrode placement**

Male participant

Female participant



The ideal electrode placement for detecting heart rate is a triangular configuration on the chest where the white and green electrodes are parallel with the heart's main axis (see illustration).

## Appendix B.10: General information sheet.

### General information

**\*Note:** Any information provided will be handled with the utmost care, and privacy is given priority at all times. None other than the researcher will handle this information, and will be safely secured at all times. The researcher in question is henceforth considered to be responsible, and thereby acknowledges his responsibility.

Please tick or fill in the boxes appropriate to your situation.

#### Age

years

#### Gender

☐ Male ☐ Female

#### Education

☐ No school completion ☐ Primary School ☐ Secondary School ☐ Vocational School  
☐ Higher Education – MSc/MA ☐ Higher Education – BSc/BA ☐ Higher Education – PhD ☐ Other

#### Profession

-----

#### Years of driving experience

years

#### Average driving engagement in the last 12 months

☐ Every day ☐ 4-6 days a week ☐ 1-3 days a week ☐ Once a month  
☐ Less than once a month ☐ Never

#### Average mileage in the last 12 months

☐ 0 ☐ 1-1,000 ☐ 1,001-5,000 ☐ 5,001-10,000  
☐ 10,001-20,000 ☐ 20,001-30,000 ☐ 30,001-50,000 ☐ 50,000+



## Appendix B.11: NASA-Task Load Index.

To be filled in by researcher:

Participant number	Task	Date and time
--------------------	------	---------------

### NASA Task Load Index

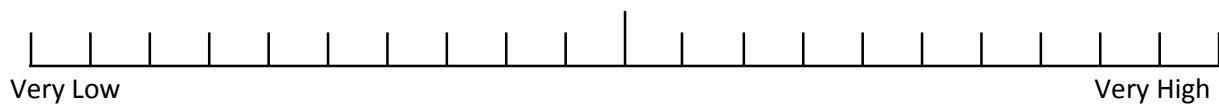
Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

#### Instructions:

Please put a cross (X) **ON THE STALKS**, to indicate the degree of the requested demands you experienced during your recent task.

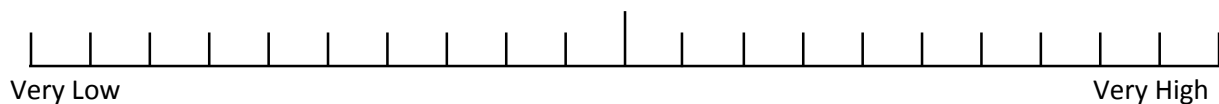
#### Mental Demand

How mentally demanding was the task?



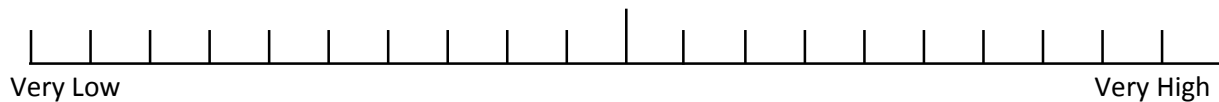
#### Physical Demand

How physically demanding was the task?



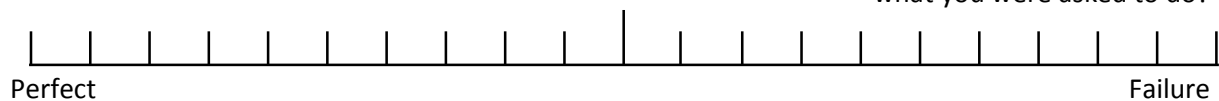
#### Temporal Demand

How hurried or rushed was the pace of the task?



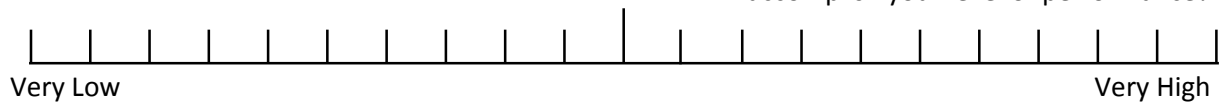
#### Performance

How successful were you in accomplishing what you were asked to do?



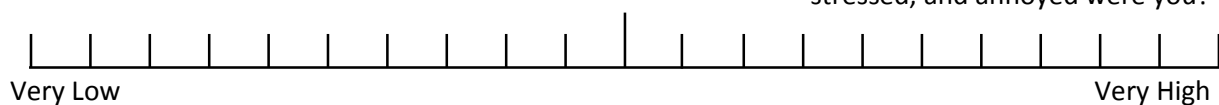
#### Effort

How hard did you have to work to accomplish your level of performance?



#### Frustration

How insecure, discouraged, irritated, stressed, and annoyed were you?



### Appendix B.12: Supplementary material: DSSQ 13 scales results, Conditionally.

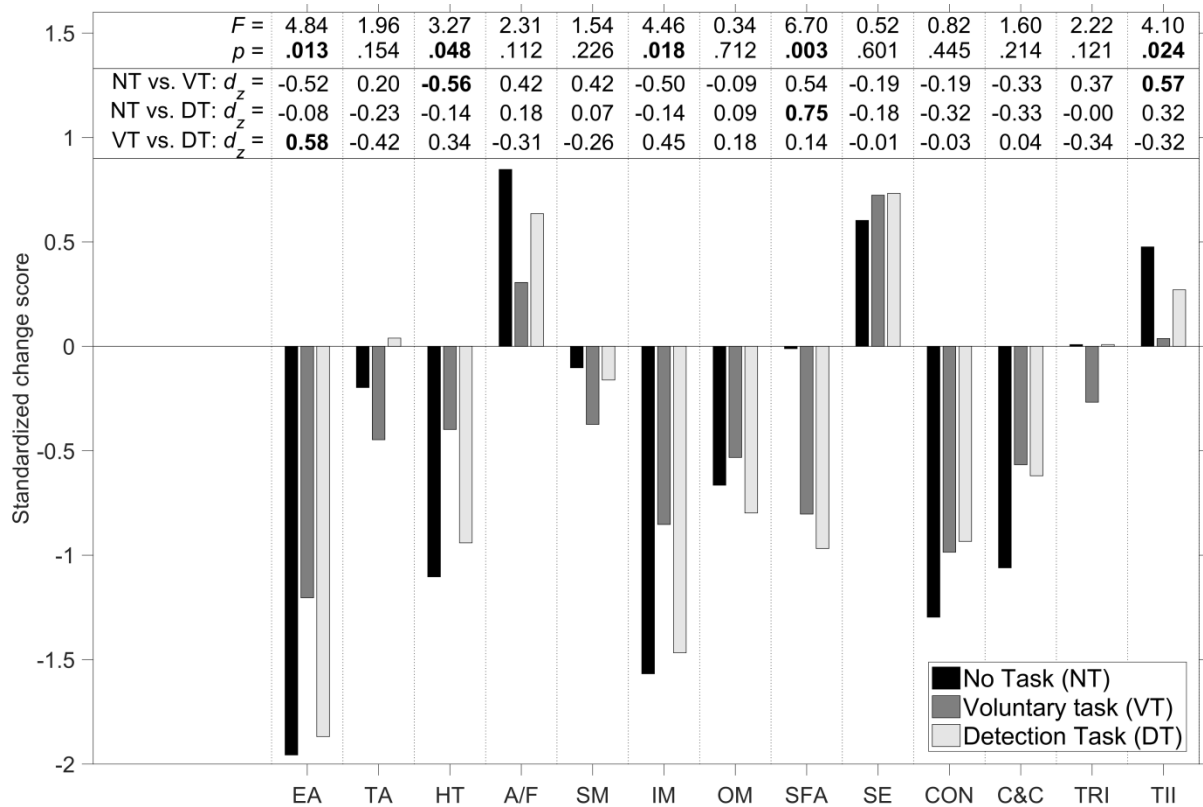


Figure S1. Standardized change scores for the DSSQ for the three experimental conditions.

EA = Energetic arousal (8 items), TA = Tense arousal (8 items), HT = Hedonic tone (8 items), A/F = Anger / frustration (5 items), SM = Success motivation (7 items), IM = Intrinsic motivation (7 items), OM = Overall motivation (1 item), SFA = Self-focused attention 8 items), SE = Self-esteem (7 items), CON = Concentration (7 items), C&C = Control and Confidence (8 items), TRI = Task-related interference (8 items), TII = Task-irrelevant interference (8 items). Also shown are the  $F$  statistic and  $p$  value for the repeated measures ANOVA and the Cohen's  $d_z$  effect size between the three pairs of conditions. For the ANOVA, the  $p$  value is shown in boldface if  $p < .05$ . For the pairwise comparisons, the  $d_z$  is shown in boldface if  $p < .05/3$ .

**Appendix B.13: Supplementary material: DSSQ 13 scales results, Chronologically.**

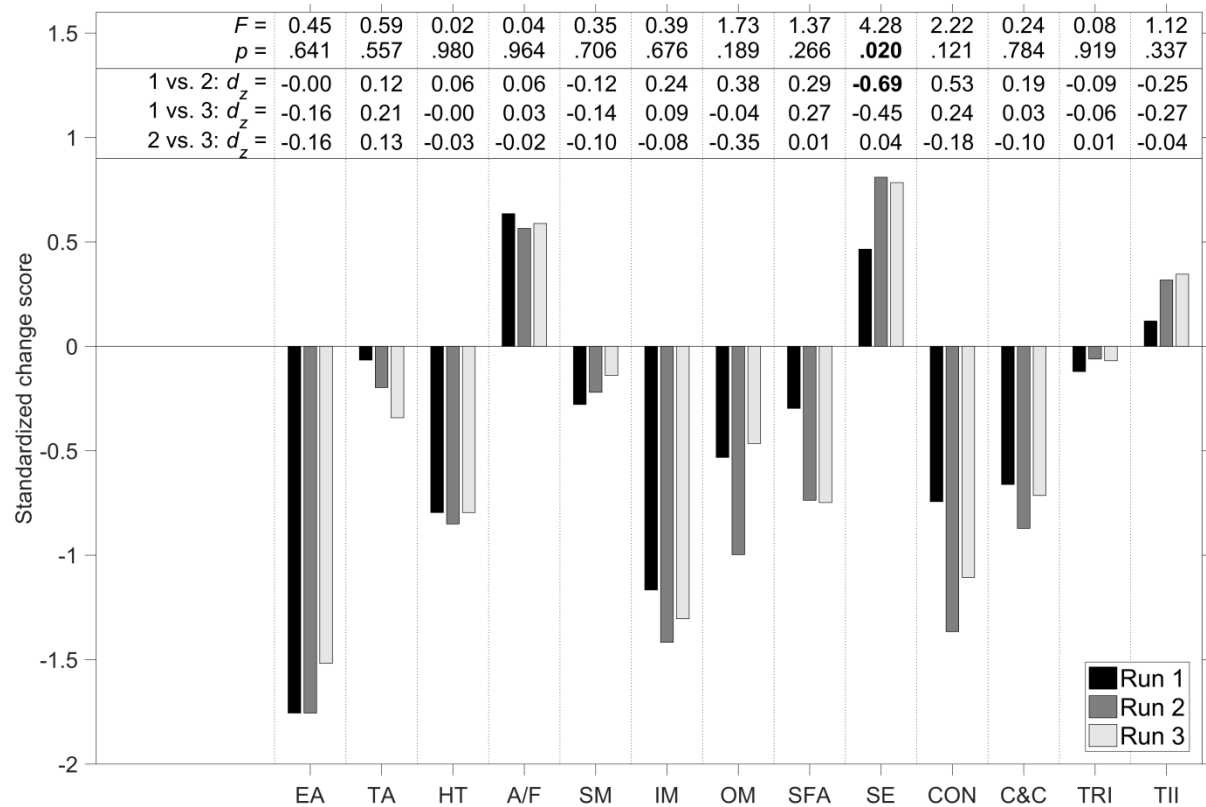


Figure S2. Standardized change scores for the DSSQ for the three experimental runs. EA = Energetic arousal (8 items), TA = Tense arousal (8 items), HT = Hedonic tone (8 items), A/F = Anger / frustration (5 items), SM = Success motivation (7 items), IM = Intrinsic motivation (7 items), OM = Overall motivation (1 item), SFA = Self-focused attention 8 items), SE = Self-esteem (7 items), CON = Concentration (7 items), C&C = Control and Confidence (8 items), TRI = Task-related interference (8 items), TII = Task-irrelevant interference (8 items). Also shown are the  $F$  statistic and  $p$  value for the repeated measures ANOVA and the Cohen's  $d_z$  effect size between the three pairs of conditions. For the ANOVA, the  $p$  value is shown in boldface if  $p < .05$ . For the pairwise comparisons, the  $d_z$  is shown in boldface if  $p < .05/3$ .

## Appendix C: Experimental Materials of Chapter 4

### Appendix C.1: 2-back letter sequence score sheet.

Utterance	Response	Utterance	Response	Utterance	Response	Utterance	Response	Utterance	Response	Utterance	Response
Q		F		B		X		B		M	
H		K		X		K		M		Q	
R		K		F		B		X		H	
X		F		H		M		K		K	
M		Q		M		R		H		M	
B		M		F		H		R		B	
F		B		Q		K		Q		X	
K		R		K		R		F		Q	
R		X		X		M		X		R	
F		H		R		H		F		F	
X		Q		B		B		K			
M		F		H		F		M			
Q		B		M		X		H			
K		H		X		Q		Q			
B		X		H		Q		B			
H		R		F		F		R			
K		K		R		K		R			
H		M		Q		H		K			
R		B		B		R		Q			
X		Q		K		X		H			
B		X		F		M		F			
M		K		B		B		X			
Q		F		Q		B		M			
F		M		X		R		B			
M		R		H		H		X			
R		H		R		X		H			
X		M		K		F		K			
H		Q		M		M		B			
Q		R		F		K		F			
B		K		Q		Q		R			

## **Appendix C.2: Experiment instructions sheet, Run 1, Condition Low Task Demands.**

### **Experiment Instructions**

First of all, thank you for your interest in this experiment. Please read the instructions carefully.

#### **Before the experiment**

You are now seated in our Southampton University Driving Simulator (SUDS). This simulator is able to simulate manual, as well as automated driving amongst a variety of scenarios.

Now you will be asked to read and fill out a few questionnaires. These are:

- Participant Information Sheet
  - o This will explain what the research is about and what can and will happen before, during and after the experiment
- Consent Form
  - o This is to ensure you understand what your participation entails
- ECG Electrode placement
  - o This is to inform you of how and where the ECG electrodes will be placed
- General Information
  - o This is to get some general information, solely for research purposes
- State Questionnaire
  - o This is to assess your current state, solely for research purposes
- Quick Association Check
  - o This is to assess your general idea on driving automation

The researcher will monitor your eye- and head movements by means of a head-mounted eye tracker. Some calibrations will have to be made, and therefore your cooperation is appreciated. Also some of your biometric measures, such as your heart rate, will be measured. For this, some electrodes have to be placed on you, as explained in the Information Sheet. Furthermore, you will be asked to “think aloud” several times along the ride. This will be rehearsed before the experiment starts, to give you an idea of what is expected from you. Lastly, during the experiment your utterances will be recorded and saved anonymously, stored and to be used by the research team only.

#### **During the experiment**

You will take part in a simulated drive in a highly automated platoon. This means you will be part of a group of highly automated vehicles that will drive very close together.

Driving in a highly automated vehicle means that the vehicle automation takes care of both the speed and the steering, so, theoretically, you don’t have to do a thing! But an incident is always possible, so you still have to pay attention. The drive will take place along a highway with mild traffic. This first drive (out of three) will be around 40 minutes.

Your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears. Try to avoid any accidents at all times, and remember to “think aloud” when asked by the researcher.

#### **After the experiment**

You will be asked to fill out three questionnaires:

- State Questionnaire
  - o This is to assess your current state after your drive
- NASA-TLX
  - o This is to assess your experiences about your drive
- Quick Association Check
  - o This is to assess your general idea on driving automation

Then you will receive the next instructions for your second drive.

### **Appendix C.3: Experiment instructions sheet, Run 2/3, Condition Low Task Demands.**

#### **Experiment Instructions**

You will now take part in the next simulated drive. This time, your only task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears.

Please remember to “thinking aloud” when requested, which will be in a similar fashion as the previous session.

After the experiment you will be asked to further develop the map and fill out the same two questionnaires again:

- State Questionnaire
- NASA-TLX

## **Appendix C.4: Experiment instructions sheet, Run 1, Condition Medium Task Demands, Part: During the experiment.**

### **During the experiment**

You will take part in a simulated drive in a highly automated platoon. This means you will be part of a group of highly automated vehicles that will drive very close together.

Driving in a highly automated vehicle means that the vehicle automation takes care of both the speed and the steering, so, theoretically, you don't have to do a thing! But an incident is always possible, so you still have to pay attention. The drive will take place along a highway with mild traffic. This first drive (out of three) will be around 40 minutes.

Your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears. Try to avoid any accidents at all times, and remember to "think aloud" when asked by the researcher.

Furthermore, you will hear a sequence of utterances of several consonants.

You are encouraged, but not required to repeat the consonant you heard 2 consonants ago.

For example:

You have heard the sequence "T, L, X, H, R".

At the point you hear the "R", you say "X", as that is the consonant uttered 2 consonants ago.

So, the proper response to this entire sequence would thus be "... , ..., T, L, X".

## **Appendix C.5: Experiment instructions sheet, Run 2/3, Condition Medium Task Demands.**

### **Experiment Instructions**

You will now take part in the next simulated drive. This time, your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears.

Please remember to “thinking aloud” when requested, which will be in a similar fashion as the previous session.

Furthermore, you will hear a sequence of utterances of several consonants.

You are encouraged, but not required to repeat the consonant you heard 2 consonants ago.

For example:

You have heard the sequence “T, L, X, H, R”.

At the point you hear the “R”, you say “X”, as that is the consonant uttered 2 consonants ago.

So, the proper response to this entire sequence would thus be “..., ..., T, L, X”.

After the experiment you will be asked to further develop the map and fill out the same two questionnaires again:

- State Questionnaire
- NASA-TLX



## **Appendix C.6: Experiment instructions sheet, Run 1, Condition High Task Demands, Part: During the experiment.**

### **During the experiment**

You will take part in a simulated drive in a highly automated platoon. This means you will be part of a group of highly automated vehicles that will drive very close together.

Driving in a highly automated vehicle means that the vehicle automation takes care of both the speed and the steering, so, theoretically, you don't have to do a thing! But an incident is always possible, so you still have to pay attention. The drive will take place along a highway with mild traffic. This first drive (out of three) will be around 40 minutes.

Your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears. Try to avoid any accidents at all times, and remember to "think aloud" when asked by the researcher.

Furthermore, you will hear a sequence of utterances of several consonants.

This time, you are to repeat the consonant you heard 2 consonants ago every time.

For example:

You have heard the sequence "T, L, X, H, R".

At the point you hear the "R", you say "X", as that is the consonant uttered 2 consonants ago.

So, the proper response to this entire sequence would thus be "... , ..., T, L, X".

## **Appendix C.7: Experiment instructions sheet, Run 2/3, Condition High Task Demands.**

### **Experiment Instructions**

You will now take part in the next simulated drive. This time, your task is to monitor the road and the drive along the ride, and intervene whenever an unexpected critical situation appears.

Please remember to “thinking aloud” when requested, which will be in a similar fashion as the previous session.

Furthermore, you will hear a sequence of utterances of several consonants.

This time, you are to repeat the consonant you heard 2 consonants ago every time.

For example:

You have heard the sequence “T, L, X, H, R”.

At the point you hear the “R”, you say “X”, as that is the consonant uttered 2 consonants ago.

So, the proper response to this entire sequence would thus be “..., ..., T, L, X”.

After the experiment you will be asked to further develop the map and fill out the same two questionnaires again:

- State Questionnaire
- NASA-TLX

## **Appendix C.8: Participant information sheet.**

### **Participant Information Sheet**

**Study Title:** Long-Term Effects of Highly Automated Platoons

**Researcher:** Daniel D. Heikoop

**Ethics number:** 18070

**Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.**

#### **What is the research about?**

This study is part of a larger international research project called HF-AUTO. I am interested in the long-term effects of highly automated driving on driving behaviour. This research is funded by the Marie Curie-Skłodowska Actions.

#### **Why have I been chosen?**

You have been approached to take part in this study as part of an effort to recruit people who have a driving license, and are at least in adequate physical and mental health.

#### **What will happen to me if I take part?**

Should you choose to take part, you will be asked to participate in three sessions of driving in the University of Southampton's Driving Simulator. The three sessions will all involve driving a similar route on a highway within a platoon of highly automated vehicles, each lasting approximately 40 minutes, with one short questionnaire to be filled out before, and two short questionnaires to be filled out after each session, plus you will be asked to create a map using A3 sheets of paper.

In total, the experiment should last around three hours. The questionnaires will ask about the amount of workload you feel you are under for each driving trial and about your feelings and thoughts at the moment. Furthermore your heart rate will be measured during the ride by electrodes placed on the upper chest- and back region, and your eye movements will be measured by means of a remote eye tracker. Both measurements will not hurt in any sense, and discretion is being considered at all times.

#### **Are there any benefits in my taking part?**

While individual benefit may be limited, your participation will help us to build an understanding of the long-term effects of highly automated driving in platoons on driving behaviour. It is hoped that the results achieved in this research will be used in the development of driving automation in commercially available road vehicles.

#### **Are there any risks involved?**

Driving in a simulator can cause motion sickness in some people, though not all. Some suggest around 10% of people experience symptoms such as nausea and dizziness. Previous experiments similar to this one resulted in no occasions where that occurred. Should this occur, participation will be stopped. Beyond the risk of motion sickness, this study does not carry with it any significant risks additional those you would experience in normal, day-to-day life.

#### **Will my participation be confidential?**

No identifying data will be released to anyone other than the main investigator. Consent forms will be stored securely in a locked cabinet, with the primary investigator having the only key.

#### **What happens if I change my mind?**

You are completely free to end your participation at any time. You are not obliged in any way to continue with the session, or even begin the session, should you so choose.

#### **What happens if something goes wrong?**

In the unlikely case of concern or complaint, please contact the Research Governance Office at the University of Southampton).

#### **Where can I get more information?**

If you would like more information, please feel free to email me, the main investigator, at D.D.Heikoop@soton.ac.uk.

**Appendix C.9: Consent form.**

**CONSENT FORM**

**Study title:** Long-Term Effects of Highly Automated Platoons

**Researcher name:** Daniel D. Heikoop

**Study reference:** -

**Ethics reference:** 18070

*Please initial the box(es) if you agree with the statement(s):*

I have read and understood the information sheet and have had the opportunity to ask questions about the study

☐

I agree to take part in this research project and agree for my data to be used for the purpose of this study

☐

I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected

☐

I understand that my participation entails the monitoring of heart rate- and eye measures, and I am happy for it to be so

☐

I understand that my participation entails the monitoring and storing of utterances, and I am happy for it to be so

☐

***Data Protection***

*I understand that information collected about me during my participation in this study will be stored on a password protected computer and that this information will only be used for the purpose of this study. All files containing any personal data will be made anonymous.*

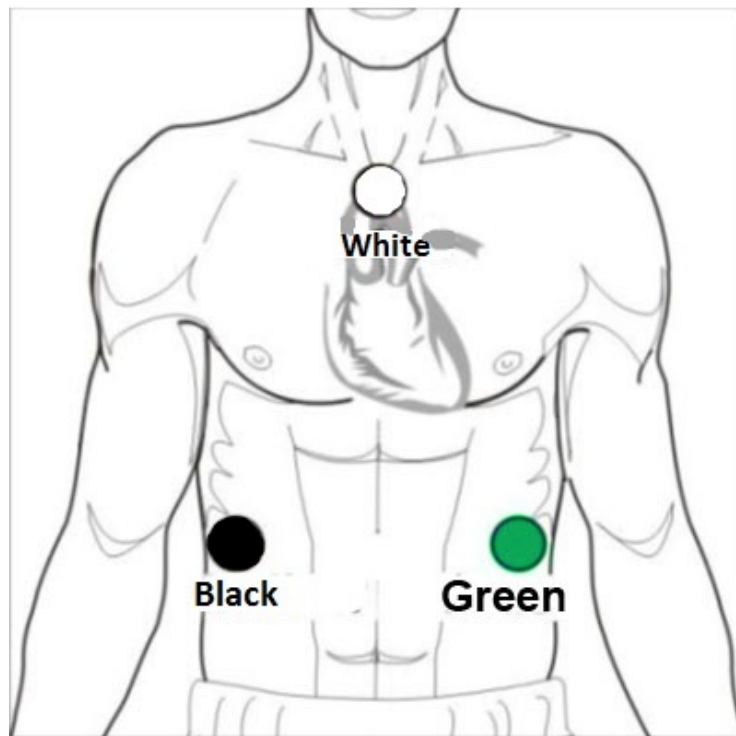
Name of participant (print name).....

Signature of participant.....

Date.....

**Appendix C.10: ECG electrode placement information sheet.**

**ECG electrode placement**



The ideal electrode placement for detecting heart rate is a triangular configuration on the chest where the white and green electrodes are parallel with the heart's main axis (see illustration).

## Appendix C.11: General information sheet.

### General information

**\*Note:** Any information provided will be handled with the utmost care, and privacy is given priority at all times. None other than the researcher will handle this information, and will be safely secured at all times. The researcher in question is henceforth considered to be responsible, and thereby acknowledges his responsibility.

Please tick or fill in the boxes appropriate to your situation.

#### Age

 years

#### Gender

☐ Male ☐ Female

#### Education

- |                                                 |                                                 |                                              |                                         |
|-------------------------------------------------|-------------------------------------------------|----------------------------------------------|-----------------------------------------|
| <input type="radio"/> No school completion      | <input type="radio"/> Primary School            | <input type="radio"/> Secondary School       | <input type="radio"/> Vocational School |
| <input type="radio"/> Higher Education – MSc/MA | <input type="radio"/> Higher Education – BSc/BA | <input type="radio"/> Higher Education - PhD | <input type="radio"/> Other             |

#### Profession

-----

#### Years of driving experience

 years

#### Average driving engagement in the last 12 months

- |                                              |                                       |                                       |                                    |
|----------------------------------------------|---------------------------------------|---------------------------------------|------------------------------------|
| <input type="radio"/> Every day              | <input type="radio"/> 4-6 days a week | <input type="radio"/> 1-3 days a week | <input type="radio"/> Once a month |
| <input type="radio"/> Less than once a month | <input type="radio"/> Never           |                                       |                                    |

#### Average mileage in the last 12 months

- |                                     |                                     |                                     |                                    |
|-------------------------------------|-------------------------------------|-------------------------------------|------------------------------------|
| <input type="radio"/> 0             | <input type="radio"/> 1-1,000       | <input type="radio"/> 1,001-5,000   | <input type="radio"/> 5,001-10,000 |
| <input type="radio"/> 10,001-20,000 | <input type="radio"/> 20,001-30,000 | <input type="radio"/> 30,001-50,000 | <input type="radio"/> 50,000+      |

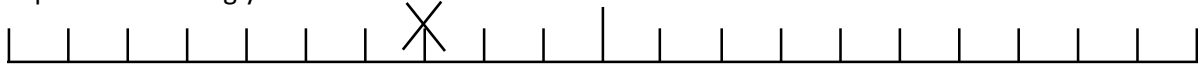
## Appendix C.12: NASA-Task Load Index.

### NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

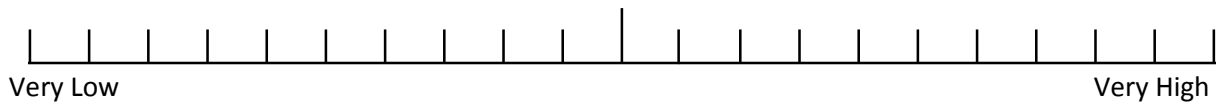
#### Instructions:

Please put a cross (X) **ON THE STALKS**, to indicate the degree of the requested demands you experienced during your recent task. Like so:



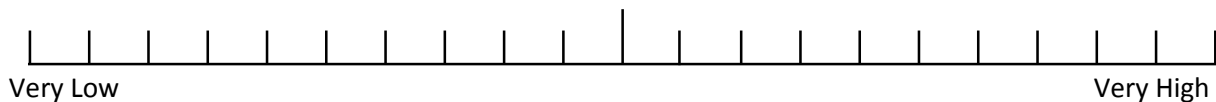
#### Mental Demand

How mentally demanding was the task?



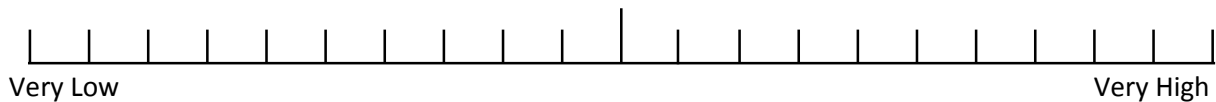
#### Physical Demand

How physically demanding was the task?



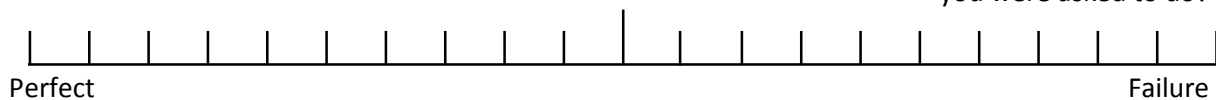
#### Temporal Demand

How hurried or rushed was the pace of the task?



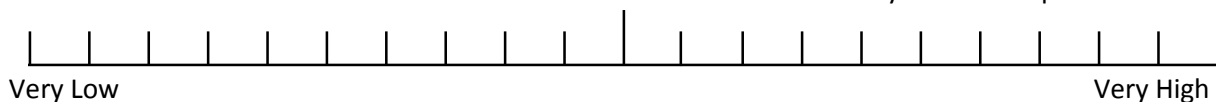
#### Performance

How successful were you in accomplishing what you were asked to do?



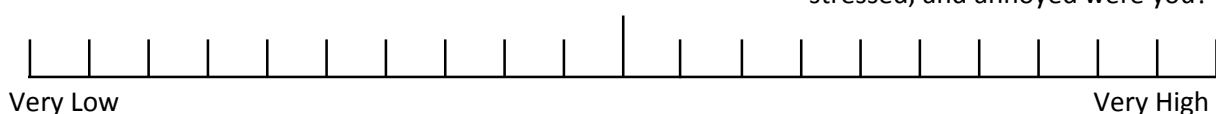
#### Effort

How hard did you have to work to accomplish your level of performance?



#### Frustration

How insecure, discouraged, irritated, stressed, and annoyed were you?



## Appendix D: Experimental Materials of Chapter 5

### Appendix D.1: Participant information sheet.<sup>1</sup>

#### Participant Information Sheet

**Study Title:** Driver Behaviour in Highly Automated Platoons

**Researcher:** Daniel D. Heikoop

**Ethics number:** 19091

**Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.**

#### **What is the research about?**

This research is part of a larger international research project called HF-AUTO, funded by the Marie Curie-Skłodowska Actions. This specific research is a collaboration between Jaguar/LandRover and the University of Southampton. I am interested in the effects of highly automated driving on driving behaviour.

#### **Why have I been chosen?**

You have been approached to take part in this study as you are the proud owner of a Tesla Model S, which features both longitudinal and lateral automated control.

#### **What will happen to me if I take part?**

Should you choose to take part, you will be asked to participate in two sessions of driving in your own vehicle along a motorway nearby. The two sessions will both involve driving a similar route on a highway within a platoon of highly automated vehicles, each lasting approximately 40 minutes, with one short questionnaire to be filled out before, and two short questionnaires to be filled out after each session.

In total, the experiment should last around two and a half hours. The questionnaires will ask about the amount of workload you feel you are under for each driving trial and about your feelings and thoughts at the moment. Furthermore, your heart rate will be measured during the ride by electrodes placed on the upper chest- and back region, and your eye movements will be measured by means of a head-mounted eye tracker. Both measurements will not hurt in any sense, and discretion is being considered at all times.

#### **Are there any benefits in my taking part?**

While individual benefit may be limited, your participation will help us to build an understanding of the behavioural effects of highly automated driving in platoons on driving behaviour. It is hoped that the results achieved in this research will be used in the development of driving automation in commercially available road vehicles.

#### **Are there any risks involved?**

A collision with the Lead Vehicle, surrounding traffic or the environment is always a possibility. However, it is expected this is not higher than during normal manual driving. In fact, the expectancy is that this chance is lower. Other than this risk, no risk is foreseen to be likely.

#### **Will my participation be confidential?**

No identifying data will be released to anyone other than the main investigator. Consent forms will be stored securely in a locked cabinet, with the primary investigator having the only key.

#### **What happens if I change my mind?**

You are completely free to end your participation at any time. You are not obliged in any way to continue with the session, or even begin the session, should you so choose.

#### **What happens if something goes wrong?**

In the unlikely case of concern or complaint, please contact the Research Governance Office at the University of Southampton.

#### **Where can I get more information?**

If you would like more information, please feel free to email me, the main investigator, at D.D.Heikoop@soton.ac.uk.

---

<sup>1</sup> Note: Although the participants actually received this information sheet prior to the experiment, none of the participants were actual owners of a Tesla. Due to time constraints it was not possible to amend the information sheet. None of the participants indicated to have a problem with this sheet.



**Appendix D.2: Participant information sheet.**

**CONSENT FORM**

**Study title:** Driver Behaviour in Highly Automated Platoons

**Researcher name:** Daniel D. Heikoop

**Study reference:** -

**Ethics reference:** 19091

*Please initial the box(es) if you agree with the statement(s):*

I have read and understood the information sheet and have had the opportunity to ask questions about the study.

☐

I agree to take part in this research project and agree for my data to be used for the purpose of this study

☐

I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected

☐

I understand that my participation entails the monitoring of heart rate- and eye measures, and I am happy for it to be so

☐

**Data Protection**

*I understand that information collected about me during my participation in this study will be stored on a password protected computer and that this information will only be used for the purpose of this study. All files containing any personal data will be made anonymous.*

Name of participant (print name).....

Signature of participant.....

Date.....

### Appendix D.3: General information sheet.

#### General information

**\*Note:** Any information provided will be handled with the utmost care, and privacy is given priority at all times. None other than the researcher will handle this information, and will be safely secured at all times. The researcher in question is henceforth considered to be responsible, and thereby acknowledges his responsibility.

Please tick or fill in the boxes appropriate to your situation.

##### Age

years

##### Gender

☐ Male ☐ Female

##### Education

☐ No school completion ☐ Primary School ☐ Secondary School ☐ Vocational School  
☐ Higher Education – MSc/MA ☐ Higher Education – BSc/BA ☐ Higher Education - PhD ☐ Other

##### Profession

-----

##### Years of driving experience

years

##### Average driving engagement in the last 12 months

☐ Every day ☐ 4-6 days a week ☐ 1-3 days a week ☐ Once a month  
☐ Less than once a month ☐ Never

##### Average mileage in the last 12 months

☐ 0 ☐ 1-1,000 ☐ 1,001-5,000 ☐ 5,001-10,000  
☐ 10,001-20,000 ☐ 20,001-30,000 ☐ 30,001-50,000 ☐ 50,000+

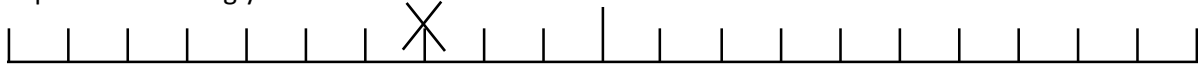
## Appendix D.4: NASA-Task Load Index.

### NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

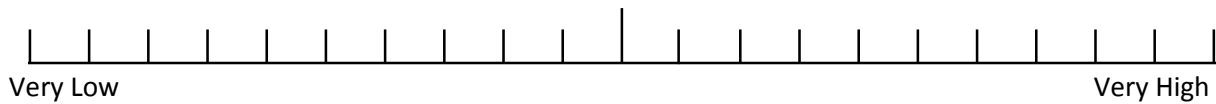
#### Instructions:

Please put a cross (X) **ON THE STALKS**, to indicate the degree of the requested demands you experienced during your recent task. Like so:



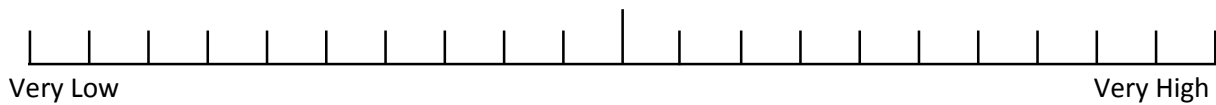
#### Mental Demand

How mentally demanding was the task?



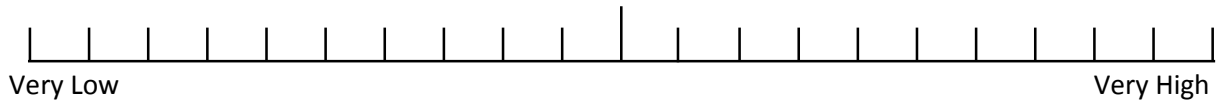
#### Physical Demand

How physically demanding was the task?



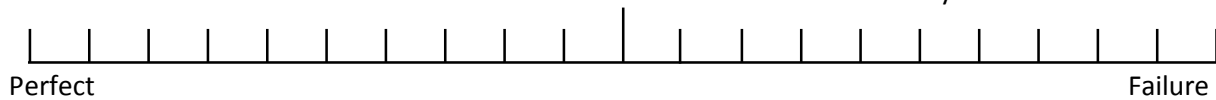
#### Temporal Demand

How hurried or rushed was the pace of the task?



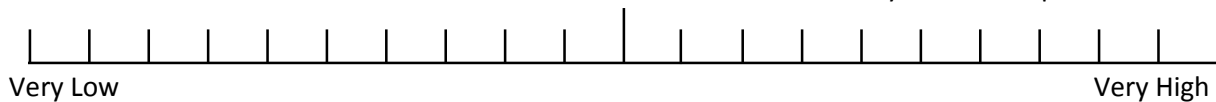
#### Performance

How successful were you in accomplishing what you were asked to do?



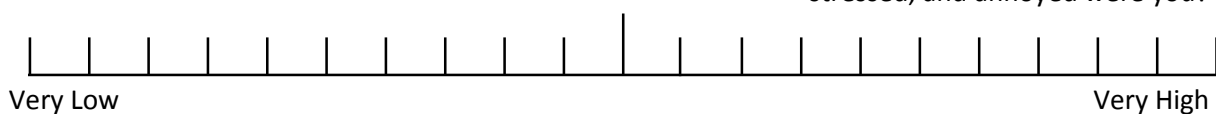
#### Effort

How hard did you have to work to accomplish your level of performance?



#### Frustration

How insecure, discouraged, irritated, stressed, and annoyed were you?



**Appendix D.5: Photographs of bridges 1-50 Northbound, retrieved from Google Maps.**



Bridge 1, NB. Location: 52.259, -1.609.



Bridge 2, NB. Location: 52.260, -1.612.



Bridge 3, NB. Location: 52.261, -1.615.



Bridge 4, NB. Location: 52.262, -1.617.



Bridge 5, NB. Location: 52.263, -1.617.



Bridge 6, NB. Location: 52.265, -1.630.



Bridge 7, NB. Location: 52.270, -1.646.



Bridge 8, NB. Location: 52.274, -1.652.



Bridge 9, NB. Location: 52.279, -1.659.



Bridge 10, NB. Location: 52.285, -1.664.





Bridge 11, NB. Location: 52.290, -1.671.



Bridge 12, NB. Location: 52.296, -1.687.



Bridge 13, NB. Location: 52.300, -1.693.



Bridge 14, NB. Location: 52.308, -1.706.



Bridge 15, NB. Location: 52.318, -1.718.



Bridge 16, NB. Location: 52.328, -1.740.



Bridge 17, NB. Location: 52.332, -1.755.



Bridge 18, NB. Location: 52.333, -1.762.



Bridge 19, NB. Location: 52.335, -1.772.



Bridge 20, NB. Location: 52.339, -1.785.



## Appendices



Bridge 21, NB. Location: 52.345, -1.794.



Bridge 22, NB. Location: 52.348, -1.806.



Bridge 23, NB. Location: 52.348, -1.812.



Bridge 24, NB. Location: 52.348, -1.826.



Bridge 25, NB. Location: 52.349, -1.837.



Bridge 26, NB. Location: 52.352, -1.858.



Bridge 27, NB. Location: 52.355, -1.885.



Bridge 28, NB. Location: 52.355, -1.887.



Bridge 29, NB. Location: 52.355, -1.898.



Bridge 30, NB. Location: 52.362, -1.948.



## Appendices



Bridge 31, NB. Location: 52.362, -1.950.



Bridge 32, NB. Location: 52.356, -1.978.



Bridge 33, NB. Location: 52.354, -1.991.



Bridge 34, NB. Location: 52.355, -2.014.



Bridge 35, NB. Location: 52.356, -2.045.



Bridge 36, NB. Location: 52.356, -2.047.



Bridge 37, NB. Location: 52.360, -2.069.



Bridge 38, NB. Location: 52.362, -2.068.



Bridge 39, NB. Location: 52.366, -2.067.



Bridge 40, NB. Location: 52.373, -2.050.





Bridge 41, NB. Location: 52.378, -2.047.



Bridge 42, NB. Location: 52.379, -2.046.



Bridge 43, NB. Location: 52.399, -2.052.



Bridge 44, NB. Location: 52.403, -2.045.



Bridge 45, NB. Location: 52.405, -2.040.



Bridge 46, NB. Location: 52.413, -2.026.



Bridge 47, NB. Location: 52.431, -2.018.



Bridge 48, NB. Location: 52.435, -2.018.



Bridge 49, NB. Location: 52.443, -2.018.



Bridge 50, NB. Location: 52.447, -2.017.



**Appendix D.6: Photographs of bridges 1-47 Southbound, retrieved from Google Maps.**



Bridge 1, SB. Location: 52.413, -2.026.



Bridge 2, SB. Location: 52.405, -2.040.



Bridge 3, SB. Location: 52.403, -2.045.



Bridge 4, SB. Location: 52.399, -2.052.



Bridge 5, SB. Location: 52.379, -2.046.



Bridge 6, SB. Location: 52.378, -2.047.



Bridge 7, SB. Location: 52.373, -2.050.



Bridge 8, SB. Location: 52.366, -2.067.



Bridge 9, SB. Location: 52.362, -2.068.



Bridge 10, SB. Location: 52.360, -2.069.



## Appendices



Bridge 11, SB. Location: 52.355, -2.068.



Bridge 12, SB. Location: 52.356, -2.047.



Bridge 13, SB. Location: 52.356, -2.045.



Bridge 14, SB. Location: 52.355, -2.014.



Bridge 15, SB. Location: 52.354, -1.991.



Bridge 16, SB. Location: 52.356, -1.978.



Bridge 17, SB. Location: 52.362, -1.950.



Bridge 18, SB. Location: 52.362, -1.948.



Bridge 19, SB. Location: 52.355, -1.898.



Bridge 20, SB. Location: 52.355, -1.887.



## Appendices



Bridge 21, SB. Location: 52.355, -1.885.



Bridge 22, SB. Location: 52.352, -1.858.



Bridge 23, SB. Location: 52.349, -1.837.



Bridge 24, SB. Location: 52.348, -1.826.



Bridge 25, SB. Location: 52.348, -1.812.



Bridge 26, SB. Location: 52.348, -1.806.



Bridge 27, SB. Location: 52.345, -1.794.



Bridge 28, SB. Location: 52.339, -1.785.



Bridge 29, SB. Location: 52.335, -1.772.



Bridge 30, SB. Location: 52.333, -1.762.



## Appendices



Bridge 31, SB. Location: 52.332, -1.755.



Bridge 32, SB. Location: 52.328, -1.740.



Bridge 33, SB. Location: 52.318, -1.718.



Bridge 34, SB. Location: 52.308, -1.706.



Bridge 35, SB. Location: 52.300, -1.693.



Bridge 36, SB. Location: 52.296, -1.687.



Bridge 37, SB. Location: 52.290, -1.671.



Bridge 38, SB. Location: 52.285, -1.664.



Bridge 39, SB. Location: 52.279, -1.659.



Bridge 40, SB. Location: 52.274, -1.652.

## Appendices



Bridge 41, SB. Location: 52.270, -1.646.



Bridge 42, SB. Location: 52.265, -1.630.



Bridge 43, SB. Location: 52.263, -1.623.



Bridge 44, SB. Location: 52.262, -1.617.



Bridge 45, SB. Location: 52.261, -1.615.



Bridge 46, SB. Location: 52.260, -1.612.



Bridge 47, SB. Location: 52.259, -1.609.