

**Let's CHAT! Check, Assess, Takeover:
Exploring the necessity and effectiveness of
behavioural interventions to guide drivers during
planned control transitions in vehicles with
conditional driving automation.**

EMILY SHAW, BSc (Hons), MSc MIEHF

**Thesis submitted to the University of Nottingham for
the degree of Doctor of Philosophy**

JANUARY 2025

Abstract

The introduction of automation technology to vehicles within the public domain promises to revolutionise everyday travel with anticipated benefits of improved road safety, comfort and mobility. Vehicles equipped with conditional driving automation, (or Level 3 (L3) automated driving systems (ADS), will be able to take control of the full driving task, allowing drivers to turn their attention towards non-driving related tasks (NDRTs) for extended periods. This automation capability is bounded by specific conditions, meaning its operation will not cover all end-to-end journeys. Therefore, drivers will be required to seamlessly switch between automated and manual driving while the vehicle is in motion following a controlled request from the automation. The introduction of vehicle automation fundamentally changes the context of the driving task overall and these planned control transitions represent a new sub-task for the human driver. However, there is a tacit assumption that experienced drivers will be able to take on any new tasks and expectations required to safely drive vehicles fitted with this capability without any additional guidance or support. The balance between the technological benefits and consumer desire brought about by the use of automated driving systems; and the maintenance of driver performance and comfort in a safety critical system is important. Therefore, this thesis proposes to test the assumption of driver competence in the use of conditional automation and consider the design and evaluation of behavioural intervention solutions aimed to encourage desired driver behaviour during planned transitions of control.

The work within this thesis details three studies (one observation study on the dataset of previous work and two evaluation driving simulator experiments) and two sections detailing the design of driver behavioural interventions to support planned control transitions. To begin, an observation study was employed to analyse the behaviour of 49 drivers who had taken part in a 5-day, longitudinal simulator study investigating control transitions in a vehicle, equipped with conditional automated driving systems. Analysis focused on driver behaviour during four planned and one emergency control transitions. Results revealed a lack of understanding from drivers on the shared control requirements during planned control transitions, and the checks and procedures necessary to safely and adequately prepare to resume manual driving. Drivers focused on improving performance in resuming operational control of the vehicle, neglected tasks associated with monitoring and planning future movements and prioritised extending shared attention with NDRTs prior to resuming manual control of the vehicle.

Findings from this foundational study were used to form an evidence-based, theoretical justification for the necessity of behavioural guidance for drivers using future vehicles equipped with conditional automated driving systems. In particular, highlighting a need to conduct more research on planned control transitions. Literature, including research

from the aviation domain, alongside the analysis of driver behaviour during the observation study inspired the concept and development of a standard operating procedure (SOP): Check, Assess, Takeover (CHAT), as a way to standardise and specify the manner (and order) that drivers should undertake tasks during a planned control transitions. This CHAT procedural framework was used to inform the design and development of two behavioural intervention solutions: the Behavioural 'CHAT' training and the CHAT Human Machine Interface. These interventions were subsequently evaluated.

A driving simulator study was designed to evaluate the effectiveness and necessity of the behavioural "CHAT" training intervention, in a between-subjects design, with 24 participants. The CHAT behavioural training intervention was evaluated against an operational training intervention based on a user manual for a commercially available vehicle fitted with automated technology. Driving performance, glance behaviour and subjective ratings of trust, workload and situational awareness were assessed. Evidence suggested that behavioural "CHAT" training had a positive influence on tactical level task performance during automated driving and following a transition to manual driving. In particular findings suggested a positive impact on visual behaviour and complying with protocols to discharge secondary devices sooner following a takeover request when compared with drivers who were given a more traditional training method (reading a user manual). Thematic analysis carried out on interviews conducted with the experienced drivers who participated in the study revealed further evidence that many drivers do not accurately understand the level of engagement with the driving task required during a planned control transition from automated to manual driving. Findings supported the view that, without adequate training or guidance, drivers find it hard to remember or, in fact, recognise the importance of attention allocation to the external driving environment during transition of control. However, responses from drivers who undertook the behavioural CHAT training suggested that it was effective in raising the need for glance behaviour that provided ongoing sampling of the external driving environment to support drivers regaining situational awareness ahead of taking over the driving task.

Finally, a driving simulator study with a between-subjects design was carried out (n=24) to investigate the effectiveness of the CHAT HMI over 3 consecutive drives when compared with a basic HMI design that gave drivers the same 60 second timeframe to prepare to drive, with audio-visual temporal cues ahead of manual driving mode engagement. Evidence demonstrated that drivers using the CHAT HMI followed the SOP, using the initial part of the transition period to discharge their devices and get themselves in position for driving, for example, checking feet position, during the first 10 seconds of the transition period. These drivers then engaged in significantly more monitoring and tracking activity engaging with the front, rear and peripheral road environment more often than the Basic HMI drivers ahead of taking

physical control of the vehicle. Monitoring behaviour was maintained more consistently across the course of the transition period by the CHAT HMI and this behaviour was consistent across the three drives. The Basic HMI group were more reactive with their glance behaviour showing temporary spikes in monitoring the external environment if something caught their attention, such as vehicles joining the motorway, but not maintaining a focus on consistent situation monitoring behaviour ahead of manual takeover. The Basic HMI group were also significantly more likely to make glances towards the HMI during the first 10 seconds after manual driving was resumed and even drivers who engaged in more rear and peripheral checks during the transition period demonstrated poorer driving performance behaviours after manual takeover, such as extended glances away from the front roadway (>2 seconds) or insufficient checks prior to changing lanes. Glance behaviour and driver performance showed that the early engagement in situation monitoring demonstrated by drivers with the CHAT HMI led to improved performance post manual driving resumption. For example, drivers who did not receive the CHAT HMI intervention were statistically more likely to be involved in a near-crash incident during a high task demand drive than those in the CHAT HMI group.

This thesis demonstrates the necessity to provide experienced drivers with additional guidance to train desired driver behaviours for the safe resumption of control of tactical as well as operational driving tasks required for planned transitions in L3 automated vehicles. The findings provide evidence that the use of a standard operating procedure to guide necessary actions in a particular order before taking over physical control of the vehicle is an effective intervention that can positively influence driver behaviour and resulting performance and comfort, during and after planned automated to manual driving transitions. The evidence provided has direct implications for the design of driver training interventions and in-vehicle HMI devices for both experienced and novice drivers of vehicles with conditional driving automation capability.

Publications

The following publications have been written within the timescale of this PhD research:

Technical Reports

Shaw, E., Large, D. R., Burnett, G. (2020). *Driver Training for Future Automated Vehicles: Introducing CHAT (CHeck, Assess, Takeover)*. London: RAC Foundation.

Conference proceedings

Shaw, E., Large, D. R., Burnett, G., (2021). Introducing CHAT to improve Driver Performance during Level 3 Automated Driving. *In Contemporary Ergonomics and Human Factors*, Warwickshire, UK.

Shaw, E., Large, D. R., Burnett, G., (2020). Towards Future Driver Training: Analysing Human Behaviour in Level 3 Automated Cars. *In Contemporary Ergonomics and Human Factors*, Warwickshire, UK.

Shaw, E., Large, D. R., Burnett, G. (2019). Ready, Set, Drive? Exploring Driver Skills and Training for Future Automated Vehicles and 'Readiness' Test Design. *In Doctoral Colloquium Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive '19)*, September 21-25, Utrecht, NL.

Journal Articles*

(*theoretical or technical contributions from PhD work i.e. simulated driving environment)

Clark, J. R., Large, D. R., Shaw, E., Nichele, E., Trigo, M. J. G., Fischer, J. E., & Stanton, N. A. (2024). Identifying interaction types and functionality for automated vehicle virtual assistants: An exploratory study using speech acts cluster analysis. *Applied Ergonomics*, 114, 104152.

Large, D. R., Harvey, C., Shaw, E., Khandeparker, S., Burnett, G., Box, E. (in press). Two for the Road: an Exploratory Study Investigating Driver and Co-passenger Interactions during Automation and the Transition of Control in a Level 3 Automated Vehicle. *Advances in Transportation Studies*

Acknowledgements

I would like to thank my supervisors, Gary Burnett, Cath Harvey and David R Large for giving me the opportunity to undertake the PhD, providing the resources to conduct my research, and ultimately, giving me the encouragement and enthusiasm to push through to the end. I owe particular thanks to David who has provided significant counsel to me over the course of what has felt like the longest roller coaster ride. Thank you for encouraging me to walk when I couldn't run, crawl when I couldn't walk and helping me to recombobulate at times of utter discombobulation! Thank you to you all, I am grateful for our fruitful discussions and your reviews of my work. Your collective support and challenge have truly helped me to grow as a researcher and develop a deep enthusiasm for all things human factors!

Thank you to the RAC Foundation for sponsoring, supporting and publishing my research on driver training for future vehicles. Thank you also to the wider Human Factors Research Group at the University of Nottingham. It has been a great group to be a part of, thank you to everyone who has provided their friendship, support and guidance through this process (as well as cake and chocolate to get us through the more challenging times!).

To my family and friends, who have stuck by me through what has felt, at times, to be a never-ending story - thank you. Your support, encouragement, dog walks, coffee breaks, animal photos, memes and packets of tissues have all been gratefully received. I promise I'm going to emerge from my office now. Finally, and most importantly, to my husband Andy and my children, Kit, Pippa and Bea - thank you for always believing in me, especially when I was struggling to believe in myself. I could not have done this without you.

Table of Acronyms

Acronym	Meaning
ACC	Adaptive Cruise Control
ADM	Automated Driving mode
ADAS	Advanced Driver Assistance system
ADS	Automated Driving System
AOI	Area of Interest
AS	Active Steering
AV	Automated Vehicle
BORIS	Behaviour Observation Research Interactive Software
BT	Behavioural Training
CDA	Conditional Driving Automation
CDM	Critical Decision Method
CHAT	Check (360°), Assess, Takeover
DDT	Dynamic Driving Task
DHW	Distance Headway
FOV	Field of View
HMI	Human Machine Interface
KPH	Kilometres per hour
KPI	Key Performance Indicator
L3	Level 3
LC	Lane Change
LCA	Lane change after boundary cross
LCB	Lane change before boundary cross
LOA	Levels of Automation
MDM	Manual Driving Mode
MPH	Miles per hour
NDRT	Non-Driving Related Task
OEDR	Object and Event Detection and Response
ODD	Operational Design Domain
OOTL	Out of the loop
OPD	Observability, Predictability, Directability
OT	Operational Training
PCT	Planned Control Transition
PCTP	Planned control transition period
PFT	Prepare For Takeover
PTD	Prepare To Drive
SA	Situational Awareness
SAE	Society for Automotive Engineers
SART	Situational Awareness Rating Technique
SAGAT	Situational Awareness Global Assessment Technique
SOP	Standard Operating Procedure
SRK	Skills Rules Knowledge
SWA	Steering Wheel Angle
TEQ	Training Evaluation Questionnaire
T(x)	Time period
TTC	Time to Collision
THW	Time Headway
TTAQ	Total Trust in Automation Questionnaire
TOR	Takeover Request
WL	Workload

Contents

Abstract	i
Publications.....	iv
Acknowledgements.....	v
Table of Acronyms	vi
List of Figures	i
List of Tables	ix
1. Introduction	1
1.1 Background	1
1.2 Research Motivation.....	3
1.3 Novel Contribution	5
1.4 Research Aims and Questions	6
1.5 Thesis Overview	8
1.6 Chapter Summaries	10
2. Literature Review	13
2.1 Introduction.....	13
2.2 Lessons from Aviation	14
2.3 The changing driving task.....	15
2.4 Visual attention and situational monitoring	17
2.5 Information processing and cognition	18
2.6 The changing needs of the human driver	20
2.7 Behavioural adaptations and procedural deviations	21
2.8 Mental models, expectations and trust	24
2.9 Driver Competence, Capability and Managing Driving Task Demands	25
2.9.1 Task Capability Interface model (Fuller, 1999)	26
2.10 Summary of literature review.....	30
3. Analysing Driver behaviour with Level 3 Automated Driving Systems	33
3.1 Introduction.....	33
3.2 Background	33
3.3 Aims and Overview	35
3.4 Method	36
3.5 Results and Discussion:	38

3.6	Conclusions and next steps.....	45
4.	Design of Driver ‘Behavioural’ Training and the ‘CHAT’ Standard Operating Procedure	47
4.1	Introduction.....	47
4.2	Behavioural change training	48
4.2.1	Learning approach	48
4.2.2	Mental model formation	51
4.2.3	The ‘CHAT’ Standard Operating Procedure.....	54
4.3	Operational training	57
4.4	Conclusions and Future Directions.....	58
5.	Driver Training Driving Simulator Study	60
5.1	Introduction.....	60
5.2	Overview and Aims	60
5.3	Methodology and approach	60
5.3.1	Experimental drive	60
5.3.2	Planned control transition	62
5.3.3	Participants	63
5.3.4	Training.....	64
5.3.5	Test Drive	65
5.3.6	Semi-Structured Interview.....	65
5.3.7	Measures	66
5.3.8	Statistical Testing.....	68
5.4	Results and Analysis	68
5.4.1	Subjective Measures.....	68
5.4.2	Visual behaviour	73
5.4.3	Driving performance.....	78
5.4.4	Non-Driving related tasks (NDRT)	78
5.4.5	Summary of Questionnaire, Behavioural and Performance Results	79
5.4.6	Thematic Analysis of Semi-structured interview	80
5.5	Discussion	97
5.5.1	Subjective ratings.....	98
5.5.2	Visual behaviour	99
5.5.3	Out-of-the-Loop, On-the-Loop and In-the-Loop	101
5.5.4	Driving performance.....	101

5.5.5	Benefits of the CHAT Standard Operating Procedure....	102
5.6	Limitations	106
5.7	Conclusions and Future Directions.....	107
6.	Design of the CHAT Human Machine Interface.....	110
6.1	Introduction.....	110
6.2	Background	110
6.3	HMI Design for Control Transitions	111
6.4	Design Approach – Conversational Interfaces	117
6.5	CHAT HMI Design.....	118
6.6	Basic HMI	124
6.7	Chapter Summary and Future Directions	125
7.	CHAT HMI Simulator Study	127
7.1	Introduction.....	127
7.2	Study Overview and Aim	127
7.3	Methodology and Approach.....	130
7.3.1	Experimental Design.....	131
7.3.2	Driving Environment and Vehicle Behaviour	132
7.3.3	Data Structure and Performance Measures.....	142
7.3.4	Participants	150
7.3.5	Human Machine Interface:.....	151
7.3.6	Procedure	152
7.4	Results	153
7.4.1	Subjective Measures.....	154
7.4.2	Driver behaviour	160
7.4.3	Glance Behaviour	162
7.4.4	Driver Capability and Control Response	190
7.5	Discussion	195
7.5.1	Subjective Situational Awareness, Workload and Trust.	195
7.5.2	Prepare for Takeover	196
7.5.3	Check, Assess, Takeover	197
7.5.4	Driver Capability and Control Response	199
7.6	Limitations	201
7.7	Conclusions.....	203
8.	Discussion and Future Directions	204
8.1	Introduction.....	204

8.1.1	Research Aims and Questions	204
8.2	Planned Control Transitions. New task: sufficient competence or necessary training?.....	205
8.3	Evaluation of design solutions	213
8.4	Overall Conclusion	219
8.5	Future Work and Design development	220
8.6	Closing Statement	224
References.....		226
Appendices.....		238
Appendix A Training Evaluation Questionnaire (TEQ)		238
Appendix B Extract from Behavioural Training – the CHAT standard operating procedure.....		239
Appendix C Operational Training User Manual		241
Appendix D Critical Decision Method (CDM) ‘Deepening’ probe questions		248
Appendix E HMI audio-visual displays for the transition from ‘Manual Driving’ to ‘Automated Driving’ mode		249
Appendix F Description and units of key data metrics recorded by the AVSimulation SCANeR software.		250
Appendix G Tables reporting Mean and Standard Deviations for the Subjective Results reported in the CHAT HMI Driving Simulator Study		251
Appendix H Tables showing Effect Sizes for Statistically Significant results from Independent T tests conducted for CHAT HMI Driving Simulator Study		252
Appendix I Summary driver performance metrics in category ‘Loss of Control = yes’		253
Appendix J Glance duration towards Roadway - descriptive statistics		254
Appendix K PCTP T-period analysis for ‘Roadway’ and a comparison between T1 and T6 for AOI in glance category ‘internal’		255
Appendix L Descriptive Statistics and Mann-Whitney Y test scores for Drive 1 lane change		257

Appendix M Table showing Mean and Standard Deviations for mirror glances at time blocks during the Drive 3 lane change.....	258
--	------------

List of Figures

Figure 1-1 Thesis Overview.....	8
Figure 2-1 Society of Automated Engineers J3016 Levels of Driving Automation (SAE International, 2021)	13
Figure 2-2: Michon's (1985) Driving skills hierarchy	15
Figure 2-3 Multi-level control in driving, including situation monitoring at each level (Merat, 2019).....	17
Figure 2-4 The features contributing to 'Task Demand', according to Fuller's (1999) Task Capability Interface model	27
Figure 2-5 The Factors contributing to driver 'Capability', according to Fuller's (1999) Task Capability Interface Model	28
Figure 2-6 The transactional relationship between Task demand and momentary driver capability (Adapted from Fuller, 2000).	29
Figure 3-1 HMI display alerts split by participant group and TOR type (Burnett et al., 2019).....	37
Figure 3-2 Driver continues to interact with laptop after 'Prepare to Drive' HMI audio TOR alert	39
Figure 3-3 Driver distracted by their mobile phone placed on the passenger seat after manual driving mode has resumed	40
Figure 3-4 Driver switching internal reading light off after manual driving mode has resumed.....	40
Figure 3-5 Driver adjusting seat during the emergency scenario TOR and audio 'resume control' alert.....	41
Figure 3-6 Driver using the countdown timer bar on the HMI display as a cue to disengage from using their NDRT	42
Figure 3-7 a) – d) presents a storyboard of driver salient behavioural adaptations from day one, day two and day five.....	44
Figure 4-1 Monitoring inherent to multi-level control in driving (Merat, 2019)	50

Figure 4-2 Training screenshot showing the SAE Levels of Vehicle Automation and the roles and responsibilities of the human driver at each level (Adapted from SAE International, 2016).....	51
Figure 4-3 Training screenshot explaining the definitions of ‘Automated’ and ‘Autonomous’ vehicles, showing blue dialogue boxes used during the presentation (Adapted from SAE International, 2016)	51
Figure 4-4 Training screenshot explaining SAE Level 3 Automation, showing blue dialogue boxes used during the presentation (Adapted from SAE International, 2016)	52
Figure 4-5 Training screenshot discussing automated driving in relation to Operational Design Domain (ODD) limitations, the role and responsibility of drivers and situation monitoring of tactical and strategic tasks to promote on-the-loop engagement	53
Figure 4-6 Training screenshot showing overview of the ‘CHAT’ standard operating procedure	55
Figure 4-7 Training screenshot of interactive slide used with expert commentary for proactive observation of the CHAT procedure. Blue ‘ego’ vehicle appears at the centre of shaded red circle	57
Figure 4-8 Page taken from the user manual created for ‘Operational training’ (full details in Appendix C)	58
Figure 5-1 University of Nottingham Human Factors Research Group Driving Simulator	61
Figure 5-2 Schematic showing experimental drive and associated HMI displays	63
Figure 5-3 Timeline of key action points to review with the driver during the post-drive interview.....	66
Figure 5-4 Example of split-screen video recordings used for analysis, showing participant engaged in NDRT during automation (top) and following the request to resume manual control (bottom)	67

Figure 5-5 Mean aggregated TEQ ratings, scaled to match original scale values. Shown for engagement, usefulness and attitude ratings with standard deviation error bars.....	69
Figure 5-6 Total Trust (TTAQ) ratings mean values with standard deviation error bars.....	70
Figure 5-7 Pre-training trust ratings showing all subscales. Mean values with standard deviation error bars.	70
Figure 5-8 Post-drive trust ratings showing all subscales. Mean values with standard deviation error bars.	71
Figure 5-9 Aggregated Situation Awareness (SART) ratings, scaled to match original rating scores. Mean values with standard deviation error bars.	71
Figure 5-10 Situational Awareness (SA) ratings showing all subscales. Mean values with standard deviation error bars.	72
Figure 5-11 Total Workload (NASA-TLX) ratings. Aggregated ratings scaled to match original scale values. Mean values with standard deviation error bars.....	72
Figure 5-12 Mean workload ratings showing all subscales. Error bars show standard deviations	73
Figure 5-13 Mean number of mirror checks made during automated driving. Includes glances to left, right and rear-view mirrors. Error bars show standard deviations	74
Figure 5-14 Mean number of mirror checks from prepare-to-drive notification (PTD) to manual driving. Includes glances to left, right and rear-view mirrors. Error bars show standard deviations	74
Figure 5-15 Number of participants who saw the tailgating, hazard car during the prepare-to-drive (PTD) period. 'Yes' indicates that participant saw the car, and vice versa	75
Figure 5-16 Initiation and duration of shared attention during prepare-to-drive (PTD) showing first driving-related glance (left end of solid bar) to	

final NDRT-related glance (right end of bar). Mean values with standard deviation error bars.....	76
Figure 5-17 Time to contemplate and undertake lane-change (LC) manoeuvre (mean values). Time zero represents the start of manual driving.....	77
Figure 5-18 example conversation between interviewer and driver [P25, OT]	86
Figure 5-19 Example 1 of driver comments regarding training before and after the review of the video of their transition of control in the simulated drive [P9, OT].	88
Figure 5-20 Example 2 of driver comments regarding training before and after the review of the video of their transition of control in the simulated drive [P21, OT].	88
Figure 7-1 Schematic of automation engagement sequence ahead of the ego vehicle crossing the trigger to initiate motorway traffic.	136
Figure 7-2 Screenshot of the residential area with working traffic lights.	137
Figure 7-3 Schematic of medium task demand scenario as the ego vehicle approaches the entry slip road of junction 29 during the PCTP.	138
Figure 7-4 image of road centre from the viewpoint of the driver of the ego vehicle at the point of MDM engagement.	139
Figure 7-5 Schematic of low task demand scenario during the PCTP. Ego and lead vehicle remain at the same speed and THW for full transition period.	139
Figure 7-6 Schematic of vehicle positions during the PCTP as the ego vehicle moves passed the entry slip road for junction 29	140
Figure 7-7 Example images taken from the driver seat, when the ego vehicle is in lane 2. Image a) shows the left-hand mirror view of lead 2 and 3 vehicles moving onto the motorway from the entry slip road at	

junction 29 (left). Image b) shows the rear-view mirror image of lead 3 in lane 3 and lead 2 in lane 2.	140
Figure 7-8 Schematic of vehicle positions at the point of MDM engagement and the movement and actions of Lead 2 (before MDM) and Lead 1 (after MDM)	141
Figure 7-9 Video split-screen image of driver engaging in a NDRT (reading a book) during ADM after switching on the in-cab reading light.	143
Figure 7-10 Images of drive glancing at (a) rear-view (b) right-wing (c) left-wing mirrors	145
Figure 7-11 Low Task demand (Drive 2) SART subscale and total SART Means and Standard Deviation (SD) error bars for HMI groups (CHAT; Basic).....	154
Figure 7-12 Medium Task Demand (Drive 1) SART subscale and total SART Means and Standard Deviation (SD) error bars for HMI groups (CHAT; Basic).....	154
Figure 7-13 High Task Demand (Drive 3) SART subscale and total SART Means and Standard Deviation (SD) error bars for HMI groups (CHAT; Basic).....	155
Figure 7-14 Low Task Demand (Drive 2) mean ratings and SD error bars for workload dimensions and cumulative 'total workload' score for each HMI condition.....	156
Figure 7-15 Medium Task Demand (Drive 1) mean ratings and SD error bars for workload dimensions and cumulative 'total workload' score for each HMI condition.....	157
Figure 7-16 High Task Demand (Drive 3) mean ratings and SD error bars for workload dimensions and 'total workload' score for each HMI condition.	157
Figure 7-17 Mean glance per minute and standard deviation error bars towards glance category 'Roadway' during the PCTP for the low, medium and high demand drives.....	163

Figure 7-18 Mean glances per minute and standard deviation error bars for Dashboard during the PCTP for the low, medium and high demand drives.....	164
Figure 7-19 Mean glances duration and standard deviation error bars for Internal AOIs during the PCTP for the low, medium and high demand drives.....	165
Figure 7-20 Mean glances per minute and standard deviation error bars for Rear/Peripheral AOIs during the PCTP for drives 1, 2 and 3.	165
Figure 7-21 Mean duration and standard deviation error bars for Rear/Peripheral AOIs during the PCTP for low, medium and high demand drives.....	166
Figure 7-22 Mean glance per minute and standard deviation error bars to 'Internal' AOIs for the low task demand drive (Drive 2) T-periods (T = 10s)	167
Figure 7-23 Mean glance per minute and standard deviation error bars to 'Rear/Peripheral' AOIs for Low Task Demand drive (Drive 2) T-periods (T = 10s)	168
Figure 7-24 Drive 1 mean glance frequency (glance/min) and standard deviation error bars for 'Internal' AOIs during each T period in the PCTP.....	170
Figure 7-25 % total glances made to internal AOI during T6 (50s – MDM).....	171
Figure 7-26 Drive 1 mean glance frequency (glance/min) and standard deviation error bars for 'Rear/Peripheral' AOIs during each T period in the PCTP.....	172
Figure 7-27 Mean glance per minute and standard deviation error bars to 'Roadway' for high task demand drive 3 T-periods (T = 10s).....	174
Figure 7-28 Mean glance per minute and standard deviation error bars to 'Internal' AOIs for high task demand drive 3 T-periods (T = 10s)...	175

Figure 7-29 Mean glance per minute and standard deviation error bars to 'Rear/Peripheral' AOIs for the high task demand drive T-periods (T = 10s)	176
Figure 7-30 Low task demand drive mean % Total Glance Duration towards key AOIs for a) T5 (40 – 50s) and b) T6 (50s – MDM).....	177
Figure 7-31 Drive 1 mean % Total Glance Duration towards key AOIs for a) T5 (40 – 50s) and b) T6 (50s – MDM).....	179
Figure 7-32 Drive 3 mean % Total Glance Duration towards key AOIs for a) T5 (40 – 50s) and b) T6 (50s – MDM).....	180
Figure 7-33 Mean glance frequency and SD Error Bars for 'Internal' AOIs during T7 (MDM – 10s).....	182
Figure 7-34 Mean Glance Frequency and SD Error Bars for 'Rear/Peripheral' AOIs during T7 (MDM – 10s)	183
Figure 7-35 Schematic showing lead 2 lane change during drive 3 ..	184
Figure 7-36 % of drivers who glanced towards the mirror during the lead 2 vehicle change lanes and % of drivers who made repeated glances towards the mirror during or following the lead 2 lane change.	184
Figure 7-37 Mean count of mirror glances with standard deviation error bars for the drive 3 lane change.	185
Figure 7-38 Example of a 'Loss of control = Yes' category.....	194
Figure 7-39 Example of a 'Loss of control = No' category	194
Figure 8-1 Mean (M) and Standard Deviation (SD) for SART dimensions for each HMI group.....	251
Figure 8-2 Mean (M) and Standard Deviation (SD) for NASA-TLX Workload dimensions for each HMI group.....	251
Figure 8-3 Mean (M) and Standard Deviation (SD) for TTAQ subscales for each HMI group.....	251
Figure 8-4 Cohen's d and effect sizes for statistically significant results of SART subgroups	252

Figure 8-5 Cohen's d and effect sizes for statistically significant results from the NASA-TLX Workload index.	252
Figure 8-6 Cohen's d and effect sizes for statistically significant results of mirror glance measures before and during the lane change in Drive 3.	252
Figure 8-7 Summary descriptive statistics of key driver performance metrics during the drive 3 lane change for 5 Basic condition drivers categorised as 'loss of control = yes.....	253
Figure 8-8 Mean glance duration and standard deviation error bars towards the roadway during the PCTP for drives 1, 2 and 3.	254
Figure 8-9 Drive 1 mean glance frequency (glance/min) and standard deviation error bars for 'Roadway' AOIs during each T period in the PCTP.....	255
Figure 8-10 CHAT HMI driver glances % AOI within internal glance category for periods T1 (TOR – 10s) and T6 (50s to MDM).	256
Figure 8-11 Descriptive statistics and Mann-Whitney U test results for Drive 1, lane change performance metrics	257
Figure 8-12 Descriptive Statistics and test scores for mirror glances 10 seconds before and during the lane change in drive 3.	258

List of Tables

Table 3-1 The number of drivers of glanced at their NDRT after the 'Prepare to Drive' TOR alert and those who glanced more than 5 times at their NDRT.....	39
Table 3-2 Number of drivers with eyes on HMI, dashboard or front windscreen at the point of manual driving mode engagement on day 1 and day 5.....	42
Table 5-1 Comparison of training approaches.....	64
Table 5-2 Summary of results from subjective measures.....	79
Table 5-3 Summary of results from objective measures.....	80
Table 6-1 HMI Display Screens for the TOR and first 30 seconds of the transition period.	120
Table 6-2 HMI Display Screens for the CHAT SOP & final 30 seconds before transition to manual driving mode.....	121
Table 6-3 Basic HMI audio-visual displays from the TOR to the resumption of Manual Driving mode (and disengagement from the automated systems)	125
Table 7-1 Experimental conditions	131
Table 7-2 Event blocks for analysis	142
Table 7-3 Glance categories and areas of interest.....	144
Table 7-4 Participant demographics by experimental condition.	151
Table 7-5 Summary of glance behaviour analysis by sub-section in 7.4.3	153
Table 7-6 Summary Results for Subjective Measures.....	159
Table 7-7 Descriptive statistics and Mann-Whitney U test results for KPIs (* depicts statistical significance $\alpha = .05$)	162
Table 7-8 Pairwise comparisons for glance duration towards 'Roadway'	166

Table 7-9 Statistically significant pairwise comparisons across T-periods (T = 10s) for low demand drive glance duration to 'Internal' ($\alpha = 0.05$).	167
Table 7-10 Pairwise comparisons for glance duration to 'Internal' categories.....	168
Table 7-11 Statistically significant simple main effects for T-periods (T = 10s) by HMI group for Drive 1 glance frequency to 'Internal' ($\alpha = 0.05$).	170
Table 7-12 Statistically significant paired comparisons across T-periods (T = 10s) for Drive 1 total glance duration to 'Rear/Peripheral' ($\alpha = 0.05$).	172
Table 7-13 Statistically significant pairwise comparisons across T-periods (T = 10s) for Drive 3 total glance duration to 'Internal' ($\alpha = 0.05$).	174
Table 7-14 Pairwise comparisons for glance duration to Internal AOIs	175
Table 7-15 Descriptive Statistics for internal glance duration for T5 and T6	175
Table 7-16 % Total Glance duration to Dashboard and HMI during T5 and T6	178
Table 7-17 % total glance duration for dashboard by HMI group and task demand.....	178
Table 7-18 Summary Results for Total PCTP	186
Table 7-19 Summary Results for PCTP T-Periods – Low Task Demand	187
Table 7-20 Summary Results for PCTP T-Periods – Medium Task Demand.....	187
Table 7-21 Summary Results for PCTP T-Periods – High Task Demand	188

Table 7-22 Key Points for T-Periods T5 (40 – 50s), T6 (50s – MDM) and T7 (MDM – 10s).....	189
Table 7-23 Key points for Events during high task demand	190
Table 8-1 Statistically significant pairwise comparisons across T-periods (T = 10s) for medium task demand total glance duration to 'Roadway' ($\alpha = 0.05$).	255
Table 8-2 Pairwise comparisons for Glance Frequency and Duration to 'Roadway' for high task demand PCTP T-periods (T = 10s).....	255

1. Introduction

1.1 Background

The introduction of automation into vehicle technology is expected to offer many benefits. These include improvements in road safety, increased mobility, enhanced driver comfort, and reductions in road congestion (Merat et al., 2012). Progression towards fully autonomous driving is steady. However, fully automated, autonomous vehicles that require no input from a human driver are unlikely to populate our roads in the foreseeable future. Until system boundaries no longer exist, the driving task is one which is shared between humans and technology (Brown and Laurier, 2017) and the human driver is required to supervise and intervene when system limits are reached (Endsley, 2017a).

The Society of Automotive Engineers (SAE) (SAE, 2016) categorises vehicle automation into 6 levels of ascending capability. Levels are classified by the extent of system intervention in vehicle control and the requirement of the human driver to monitor system performance and resume control of the driving task where necessary. Level 2 vehicles, available in the current market, provide functionality that automates lateral and longitudinal vehicle controls. However, the human driver remains responsible for monitoring the system and must be ready to take over the driving task if required. At level 3, automation capability extends to the monitoring task, where under limited and specified operating conditions, the vehicle systems can take on responsibility for object and event detection and response (OEDR). This additional capability allows drivers to switch their attention towards non-driving related tasks (NDRTs). However, as system capabilities are bounded, the human driver remains responsible for the vehicle's actions and must be ready to intervene in the event of a system failure or where a boundary limitation is reached (White et al., 2019).

Driving automation technology is developing at pace, with car manufacturers appearing to be in an autonomous driving systems arms race to make fully self-driving vehicles a reality for public consumption (Croutch, 2023). There is an expectation that we will continue to see a greater number of vehicles offering mid-levels of driving automation capability and begin to see vehicles that possess the ability to operate autonomously under certain conditions (Kyriakidis et al., 2019). Vehicles with 'partial' (level 2) or 'conditional' (level 3) driving automation, that share aspects of the driving task between the human driver and the vehicle, are likely to retain key elements of current vehicles. For example, they will look the same and provide the same primary input controls (e.g. steering wheel, foot pedals etc.). As such, these vehicles may not present an obvious step change in technological

development to drivers. However, these new vehicles represent a radical change in ideology, redefining the role of, and expectations placed upon, the driver (Sullivan et al., 2016).

The new forms of driving afforded by partial and conditional driving automation places new demands on drivers. This is because the nature of the driver's interaction with the vehicle is fundamentally changing from active control in current, manually driven vehicles, to passive monitoring, supervisory control and selective intervention. A key task within manual driving impacted by the introduction of automation is situation monitoring; this is the continual processing of perceptual information in order to understand and make predictions about the dynamic driving environment to aid decision making and control the vehicle. As more decision-making functions for vehicular control inputs are automated the driver naturally pays less attention to the driving task (Banks and Stanton, 2016, 2019). Consequently, the engagement of automated driving systems (ADS) takes drivers 'out of the loop' of control (Parasuraman et al., 2000), reducing their level of perception and comprehension of the system state and driving environment and the projection of their future state, a construct termed 'situation awareness' (Endsley, 2017a).

Drivers of vehicles with mid-levels of automation capability maintain fallback responsibility for the driving task, requiring either to respond rapidly in the event of a system failure or responding to a request to intervene, more likely as the result of a system capability boundary being approached. At level 2, where the driver maintains responsibility of the monitoring task, but no longer has physical control of the operational sub-tasks, performance issues relating to these control transitions are typically caused by driver distraction (i.e. the driver turning their attention away from passively monitoring the system and the driving environment). At level 3, the system design permits the driver to turn their attention to non-driving related tasks under certain circumstances. However, as system capabilities are bounded, the driver will need to be able to smoothly transition to full control of the driving task if, or more likely, when the boundaries of the operational design domain are reached. Therefore, drivers need to understand the level of attention and situation awareness they need to have in relation to the mode of automation and be able to calibrate these in a timely manner during dynamic operations. Nonetheless, research investigating these transitions from automated to manual control states (e.g. Dogan, et al, 2017) have highlighted performance challenges associated with the re-engagement of drivers' cognitive and perceptual-motor controls necessary to effectively take-over the driving task, suggesting that if or when a system issues a take-over request (TOR), the driver may not be ready to drive.

1.2 Research Motivation

This new driver role, and the skills required to effectively and safely carry out the driving task in future automated vehicles, has not yet been mapped out (Banks and Stanton, 2016). Current assumptions regarding training and instructional guidance for this changing driving task is that the more functions are automated the less there is for the driver to do and therefore, little necessity for any form of intervention (Cummings et al., 2019). For experienced drivers this extends to a belief that the skills and experience gained from manual driving training is directly transferable to the automated driving context. Therefore, drivers of vehicles equipped with these advancing levels of automation receive little to no direct training on how to use them (Casner and Hutchins, 2019), with current, passive methods of instruction, such as driver manuals or dealership guidance, being used as a sufficient way to bridge any gap.

At best, these assumptions are based on reasonable assertions of driver competence, given that the introduction of other lower-level automated features, such as cruise control, have not required more instructional guidance than what is being offered here. At worst, they are formed from capitalistic motivations to get these vehicles to market as quickly as possible, playing down the human factors implications associated with the introduction of these technologies. Regardless, of the reasoning, there is mounting evidence that the assumptions around lack of necessity for additional driver training or guidance do not hold true at these mid-levels of automation and their more fundamental impact on the driving task. Firstly, at the level of conditional driving automation, the driver is almost completely decoupled from the driving task, being removed from both physical and mental control loops. Presuming that drivers will be able to smoothly and rapidly resume manual driving and action the control responses required to meet the demands of the driving task without challenge is not supported by human factors or psychological research within this area (e.g. Mole et al., 2019; De Winter et al., 2014; Endsley, 2017a). The degradation of drivers' attention allocation to the critical information in the roadway and, therefore, situational awareness. Albeit, it is also believed by a number of researchers that given the right guidance and sufficient time, levels of attention and what is known as situational awareness can be rebuilt (Endsley, 2017a; Carsten and Martens, 2019; Trosterer, et al., 2017), thereby ensuring the driver capability available from the time the takeover occurs is at an appropriate level to take forward the driving task and meet whatever dynamic demands the driver is faced with.

Secondly, it has been shown that many drivers do not accurately understand the capabilities and characteristics of the automated systems they are using or their required level of control and co-operation within the driving task (Casner and Hutchins, 2019;

McDonald, et al., 2018). Clarity and transparency of the role, responsibilities and skills required from the human driver is crucial for successful human-automation interaction. A mismatch of expectations regarding the key supervisory requirements of the human driver can have dire consequences for performance and safety (Casner and Hutchins, 2019), as already demonstrated with fatal crashes directly involving the interaction between the human driver and their vehicle's driving automation capabilities (e.g. NTSB, 2017). In fact, there is increasing evidence of unfavourable behavioural adaptations from drivers experiencing current SAE level 2 automated driving features on the road. Specifically, such drivers can be seen relying on the technology in situations for which it was not designed or intended and changing their own behaviour as a consequence of its use (Brown and Laurier, 2017; Banks et al., 2018; Lin et al., 2018). Rather than concluding that these drivers are actively choosing to *disregard* their role and responsibilities, it is proposed that such behaviour is indicative of drivers *not knowing* or *understanding* their role and responsibilities. This may be because they lack an awareness of the limits of the technology's capability. Consequently, it is proffered that there is a fundamental need to construct ways to provide drivers with opportunities to access the education, guidance and/or experiences to plug the gap in driver competence opened up with the introduction of this level of driving automation. A core proposal of the thesis is that there is an urgent requirement to investigate effective and pragmatic approaches to behavioural interventions to moderate driver expectations and guide subsequent behaviour within the new driving task, to prepare them for the uptake of, yet undefined, new skills.

This research focuses on level 3 (L3) automated vehicles (AVs), or 'Conditional Driving Automation'. As previously mentioned, L3 driving automation capability allows the driver to carry out non-driving related tasks (NDRTs) during dynamic operations, taking them out of the loop of control by design. However, the driver remains responsible for the vehicle's actions and needs to be ready to take over control of the vehicle either in the event of a system failure or where the system's capability boundary (its 'Operational Design Domain (ODD)') has been reached. Under both circumstances a takeover request (TOR) will be issued to the driver by the system and the role of the driver will change depending on who has control of the driving task. Drivers will need to build up the requisite skill set to effectively interact with the automated system and smoothly transition between automated and manual driving modes (Gold et al., 2018)

There is an abundance of research investigating transitions of control in level 2 vehicles focused on the impact of issues relating to complacency, driver distraction and vigilance decrement on takeover response times and performance. However, by the nature of the level of

automation capability, at level 2 control transitions between automated and manual driving are time critical and occur in unexpected and 'emergency' circumstances, where there has been some form of failure or capability gap of the automated systems. By contrast, at level 3 automation, expectations are that this advancement in the technology will mean that 'emergency takeover' requirements due to, for example, system failures or safety critical situations where the driver needs to intervene, will be reduced. Instead, control transitions at this level will be planned, allowing for the possibility to give drivers sufficient time to ready themselves to take over the driving task. As the automation capability defined at this level is bounded, the expectation is that most, if not all, journeys in these vehicles will require at least one planned transition of control. Therefore, this controlled or planned takeover will form a new part of the driving task in the automated context.

This marks the starting point for the thesis. It is proposed that there is a need to design training or instructional guidelines to help drivers to understand how to interact with conditional driving automation systems in order to smoothly and effectively transition between control loops. It is argued that, in the context of a vehicle equipped with level 3 automated driving system (L3 ADS), this should include sequencing tasks to get the driver 'on the loop' of control before the driver takes physical control of the vehicle. Current vehicle manufacturers supported by government policy strategy are driving momentum to get these vehicles on public roads in the near future (Kinnear et al., 2021). Therefore, my research focuses in on experienced drivers who will likely be the first to access these vehicles and beyond the point of formal training requirements (in their current form). Knowing that it is unlikely to be a desirable option to need to introduce graduated licences and further formal training to this population, any behavioural intervention design needs to be flexible enough to be able to be applied in a variety of circumstances. Ideally any instructional guidance should be designed in a way that would allow for consistent application to both experienced and novice drivers. In-vehicle HMI systems are a standard feature in modern vehicle design and offer a potential way of delivering behavioural guidance or instruction to drivers, much like the functionality of a primary flight display provides to a pilot. In consideration of these points, the approach taken in the construction of this body of work was to first investigate training solutions in a more traditional sense and then to investigate options to apply these design strategies using a technical solution in the form of an in-vehicle HMI display system.

1.3 Novel Contribution

At the time of writing this thesis, there are very few, if any, investigations into potential training or instructional solutions for drivers of vehicles equipped with L3 capability (conditional driving automation)

(beyond the passive solutions mentioned previously). Certainly, none who have taken the same approach as described here. There is still an abundance of research studies focused on the time it takes to takeover from the automated systems, even with the reduction in criticality at L3. However, as per the recent challenge put forward to broaden thinking outside of this paradigm (De Winter et al., 2021), it is pondered if focusing on how quickly drivers can take over and perform well in that momentary transfer of control is the wrong question in answer to the wrong problem in the context of L3 automation capability. Instead, it is proffered that the issue faced is more around ensuring drivers have the available capability required to match the moment-to-moment dynamic demands of the driving task once they resume manual control. Knowing that both driver competence via education, training and experience, and human factors issues, such as reduced situational awareness contribute to available driver capability (Fuller, 2000), the question becomes how can drivers be guided and supported to rebuild the situational awareness they need to reach a level of driver capability to satisfy the requirements of the driving task beyond the transition of control to manual driving? A key part of both this problem statement and subsequent question is whether this is, in fact, a necessary concern. As previously highlighted, assumptions are that experienced drivers will already possess the skills to overcome any such issues with control transitions and will quickly adapt to performing the driving task in the automated context. Therefore, the starting point of the thesis is to explore the ways experienced drivers interact with these vehicles of the future to test the necessity for behavioural interventions and identify any driver behavioural adaptations that can help to guide design solutions.

1.4 Research Aims and Questions

The overall aim of the thesis is to investigate design solutions to encourage desired driver behaviour and interaction with vehicles capable of Conditional Driving Automation (SAE level 3 vehicles). It is the working assumption of this research that as level 3 automation capability allows for object and event detection and response (OEDR), the key requirement relating to automated to manual transitions of control will be the issue of a takeover request from the automated system for a planned transition of control due to an operational design domain boundary rather than an emergency or critical context. Therefore, it is these planned control transitions that set the context for its investigations. The body of work within this thesis was split into 3 research aims, which were used to shape and develop the research over time. A key objective of this research is to explore the necessity of interventions for experienced drivers (i.e. drivers who hold a current UK drivers' licence with a minimum of 3 years driving experience) to identify key behaviours and actions required to assume the redesigned driver role, in particular, be able to safely and effectively resume the driving

task from the automated systems under controlled conditions (i.e. a planned transition scenario) during dynamic operations.

This research focused on the human factors challenges that result from drivers being able to disengage from the dynamic driving task (DDT) and turn their attention to NDRTs resulting in an out of the loop state. However, it is acknowledged that the removal of the human from the loop of control is by design and there is an important balance to be struck between the technological benefits and consumer desire brought about by the ability to engage with NDRTs during automated driving, and the maintenance of driver performance and comfort in a safety critical system. Therefore, a further objective of this research was to investigate the effectiveness of design solutions that aimed to: support the formation of appropriate mental models of the new driving task; and encourage behaviour aimed at inducing an 'on-the-loop' state in drivers from the point a controlled takeover request is issued during automated driving to the point where manual driving control is resumed.

To summarise, the 3 research aims of this thesis are as follows:

1. To identify human factors issues relating to driver behaviour in vehicles with conditional driving automation.
2. To investigate driver interactions and behavioural adaptations in the context of L3 'Conditional Driving Automation' and planned control transitions.
3. To design and evaluate the effectiveness of training and technological solutions to support desired driver behaviour during planned control transitions in vehicles with conditional driving automation.

Research questions

There are four key research questions driving my research:

1. RQ1. What can we learn from how drivers interact with L3 AVs to inform driver training and behavioural requirements?
2. RQ2. What behaviours and actions are necessary and sufficient to prepare drivers to resume the dynamic driving task (DDT) from automated driving during dynamic operations?
3. RQ3. How can driver training support the uptake and maintenance of desired driver behaviour and provide a countermeasure to undesirable behavioural adaptations/operational deviations?
4. RQ4. How can technological solutions (e.g. in-vehicle HMI display systems) support an 'on the loop' state in drivers during automated driving to support the calibration of SA and attention for smooth transitions of control?

1.5 Thesis Overview

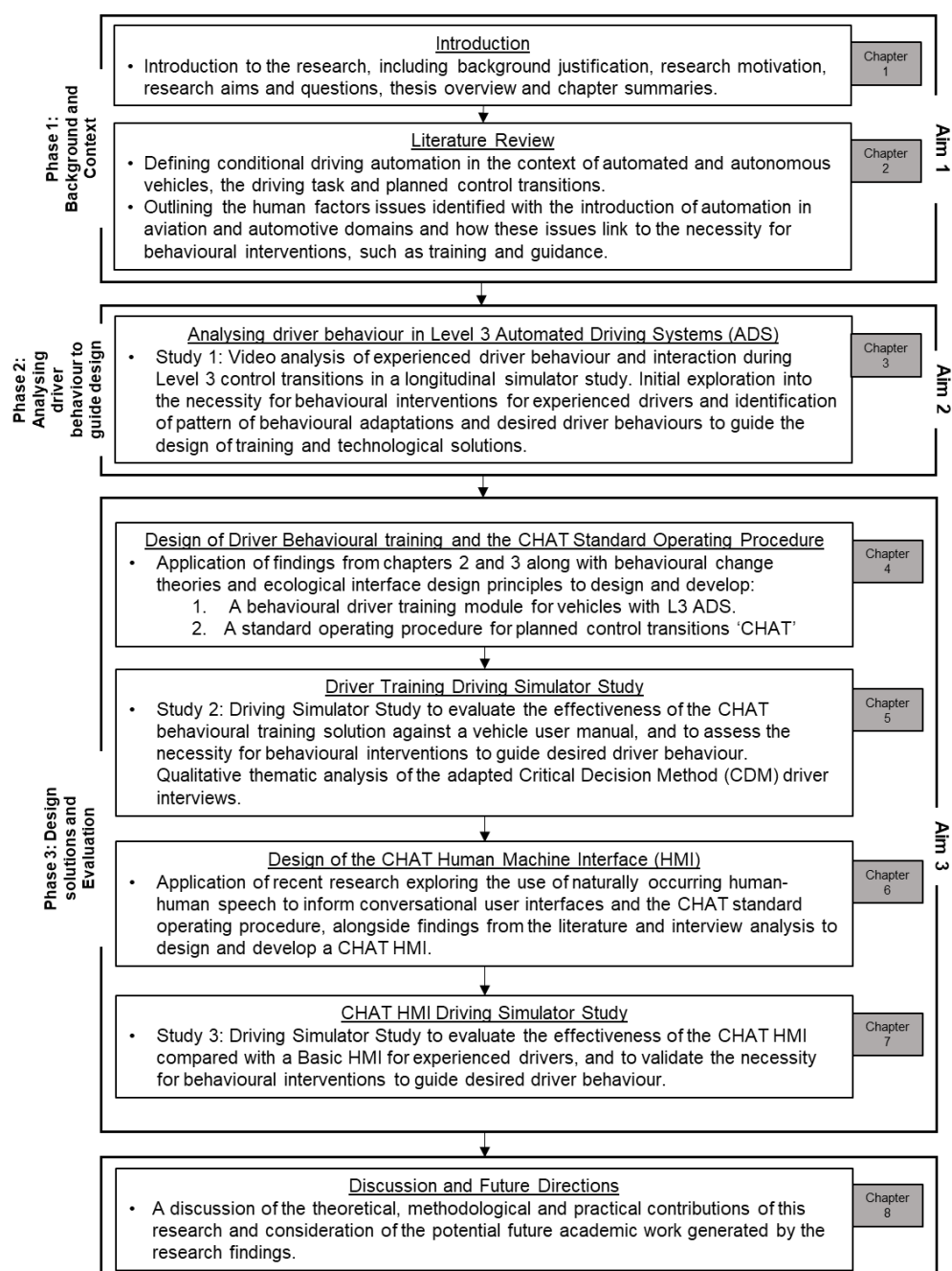


Figure 1-1 Thesis Overview

An overview of the thesis is shown in **Error! Reference source not found.** This overview shows how each phase of the thesis was applied to respond to the core research aims and how each section led to the cumulative development of ideas through the body of work. The thesis combines a mixture of explorative methods and evaluative techniques to address the core research questions. Employing a mixed methods approach drew together a broad spectrum of data sources, ranging from key research within the field of automation and human factors to analysis of objective performance data and subjective views of

experienced drivers themselves. Methods were chosen as being the most appropriate for the aims of both the phase of research and specific studies. Methods employed during phases one and two focused on data gathering and exploration to provide a starting point for investigation. This included an extensive review of the literature, alongside an observation study that involved analysis of driver behaviour drawn from extensive video data recorded during a previous longitudinal driving simulator study focused on conditional driving automation that was conducted by the research team at the University of Nottingham. The original study (see Burnett et al., 2019; Large et al., 2019) was not conducted as part of this thesis. However, the additional behavioural analysis conducted on the vast volumes of driver recording was formed and produced from the work in this thesis. The outputs of this additional observation study contributed to the results and conclusion for the study by Burnett et al. (2019). The full study and findings associated with this thesis are reported as part of this body of work. Overall, the combined findings drawn from these early phases of work provided a baseline for the thesis by informing two key outputs. Firstly, an evidence-based, theoretical justification for the necessity for behavioural interventions to support and guide drivers during planned control transitions. Secondly, the design and development of a standard operating procedure for the task of planned control transition and behavioural training solution.

Phase three involved the design and development of two proposed design solutions, in the form of behavioural training and an HMI display system, tied together by the standard operating procedure. These interventions were each evaluated using a separate, comprehensive driving simulator study. The use of a driving simulator set up to emulate use of a vehicle with conditional driving automation allowed a safe, cost effective and controlled assessment of human interaction with technology that is not yet readily available on public roads. The findings from which can provide evidence to support a proof of concept for these potential intervention solutions. In addition, a qualitative cognitive task analysis interview method was adapted and employed during the first simulator study used to evaluate the behavioural training intervention. Use of this adaptation of the 'critical decision method' interview provided rich data for a thematic analysis to be carried out on the views and behaviour of the experienced drivers involved in the study. An extended review of the literature combined with the findings from the thematic analysis allowed for a reflective evaluation of the work conducted in the thesis so far, including the use of the proposed standard operating procedure and training intervention. These findings could then be used to inform the design and development of both the technological intervention solution and the experimental design of the final simulator study.

1.6 Chapter Summaries

Chapter one introduces the thesis. This initial chapter introduces the area of driving automation, the new control transition driving subtask and the research motivation to investigate driver training for vehicles equipped with conditional driving automation or, 'level 3' systems. It outlines the research aims and questions and provides a summary of each chapter as well as a schematic overview of the thesis.

Chapter two presents an extensive review of the literature. The thesis has a strong theoretical grounding in our understanding of the cognitive capabilities and biases of the human driver and draws in part on experience and literature from the aviation domain. It highlights the unintended consequences, and potentially deleterious effects, of introducing vehicles with conditional driving automation onto the roads without considering the needs and capabilities of the attending human driver. In particular, the review emphasises the importance of providing, from the outset, clear and consistent learning strategies to foster the development of accurate mental models that structure and reinforce the expectations and shared responsibility for the driving task with the automated driving systems (ADS).

Chapter three describes an observation study analysing over 120 hours of driving captured during a longitudinal driving simulator study carried out by the Human Factors Research Group at the University of Nottingham. The study examines the interactions of experienced drivers with vehicles fitted with level 3 automation by analysing driver behaviour recorded during 'commute style' simulated drives. Four of the five drives experienced by participants included extended periods of automated driving coupled with planned control transitions and an additional drive involved an emergency control transition. The study identifies a pattern of behaviour that sees drivers prioritise the continued use of secondary devices even after they have been alerted to request to intervene from the automated systems. In particular, it shows that drivers tend to wait until they have transitioned back to manual control before focusing on rebuilding their situational awareness. The findings from this study set the foundation for the standard operating procedure design used as the basis for the behavioural interventions developed in subsequent chapters.

Chapter four presents an outline of the first driver training intervention solution design, the behavioural 'CHAT' training. The chapter describes the development of a behavioural change training intervention primarily focused on forming adequate and correct mental models of automation; raising awareness and understanding of the impact of vehicle automation on driver behaviour and the consequences associated with being out-of-the-loop (OOTL) of control; and establishing a standardised operating procedure to specify and communicate a clear sequence of tasks required by human drivers to complete a planned

control transition. This chapter see the first introduction of the 'CHAT' (Check, Assess, Takeover) standard operating procedure and provides a comprehensive explanation about its conception and development, including the semantic representation of communication between team members during joint operations.

Chapter five describes the first of two driving simulator-based experimental studies. The chapter describes the evaluation study of the 'CHAT' behavioural training intervention when compared with a typical training method offered for vehicles with additional automation, in the form of an operating manual. Twenty-four drivers were given 15 minutes of training either using the behavioural or operational format. The drivers were tasked with driving a vehicle equipped with level 3 automation on a UK dual carriageway and given free rein to use their own secondary devices as they saw fit, during 10 minutes of automated driving. The study investigated how drivers behaved during a planned control transition and then performed in the driving task during and following the resumption of manual control. The chapter also presents the thematic analysis from the interviews conducted with the drivers. The interviews used an adapted version of the critical decision method to elicit the retrospective perceptions, expectations, goals, and judgements of the drivers about key moments during the simulated drive and the training they had undertaken. Semi-structured probing questions were used with the drivers while watching back the video from the experimental drive to aid recall.

Chapter six describes the design and development of the 'CHAT' HMI. It presents the use of an in-vehicle HMI display solution as a potential way to guide desired driver behaviour and provide a novel and effective countermeasure to the out of the loop performance problems seen during planned control transitions. It outlines the use of innovative design ideas from the research into the use of conversational user interfaces in the driving context. The design uses recent research that identified patterns from naturally occurring human-to-human speech data to guide the design of 'how' to communicate information to drivers via in-vehicle HMI displays. Key ideas from this research are used alongside the CHAT, check, assess, takeover procedure, as a standardised checklist of actions guiding 'what' information to communicate to drivers. This chapter highlights the flexibility of design provided by the CHAT procedure as well as the potential for standardisation and consistency that comes with a standard operating procedure, typically seen in the aviation domain, but also reminiscent of the mirror, signal, manoeuvre (position, speed, look) routine ubiquitous with driving lane change task guidance (DVSA, 2024).

Chapter seven describes the final driving simulator experimental study. This study evaluates the effectiveness of the CHAT HMI in comparison to a Basic HMI design that provides an initial takeover request, followed by temporal information counting down to the resumption of manual control. Twenty-four drivers took part in three experimental drives

framed as commutes. During each drive, they received a takeover request after a ten-minute period of automated driving and were required to conduct a planned control transition using the same HMI design (either CHAT or Basic) for all three drives. The study investigated how drivers behaved during a planned control transition and performed in the driving task during and following the resumption of manual control. Analysis was particularly focused on driver attentional allocation and their development of situational awareness during the transition period and how that impact subsequent driver performance following resumption of manual driving. Extensive work was carried out to design and develop the final simulator study to reflect the research and theoretical learnings from the thesis to date. Due to the significance of this work, this chapter presents a detailed justification and description of the experimental design and methodological approach. This includes the design of methods relating to measuring driver situational awareness during and following planned control transitions and identifying patterns and links to an evaluation of driver performance.

Chapter eight discusses the overall research and proposes potential future work and design proposals based on the research findings. The chapter amalgamates the various findings and knowledge developed throughout the research, highlighting how the research aims have been achieved.

2. Literature Review

2.1 Introduction

The development of automated vehicles will fundamentally change the role of the 'driver' (Sullivan et al., 2016). Progression towards higher levels of vehicle automation is in motion (Casner et al., 2016; Brown and Laurier, 2017; Kyriakidis et al., 2019). However, until system boundaries no longer exist, the driving task is one which is shared between humans and technology (Brown and Laurier, 2017), and the human driver is required to supervise and intervene when system limits are reached (Victor et al., 2018).

The SAE (2016) categorises vehicle automation into 6 levels of ascending capability. Levels are classified by the extent of system intervention in vehicle control and requirement for the human driver to monitor system performance and resume control of the driving task. SAE level 2 vehicles, available in the current market, provide functionality that automates lateral and longitudinal vehicle controls. However, the human driver remains responsible for monitoring the system and must be ready to take over the driving task when required. Many drivers do not accurately understand the capabilities and characteristics of these level 2 automated systems they are using, or the required level of control and engagement with the driving task (McDonald et al., 2018; Casner and Hutchins, 2019). Clarity and transparency of the role, responsibilities and skills required from the human driver are crucial for successful human-automation interaction. Dislocation of expectations of the key supervisory requirements of the human driver can have dire consequences for performance and safety (Casner and Hutchins, 2019), as already demonstrated with the first fatal crash of the Tesla Model S vehicle in May 2016 (National Transportation Safety Board, 2017).

SAE LEVEL 0™		SAE LEVEL 1™		SAE LEVEL 2™		SAE LEVEL 3™		SAE LEVEL 4™		SAE LEVEL 5™							
You are driving whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering						You are not driving when these automated driving features are engaged – even if you are seated in “the driver’s seat”											
You must constantly supervise these support features; you must steer, brake or accelerate as needed to maintain safety						When the feature requests, you must drive		These automated driving features will not require you to take over driving									
Copyright © 2021 SAE International.																	
These are driver support features						These are automated driving features											
What do these features do?						These features are limited to providing warnings and momentary assistance		These features provide steering OR brake/acceleration support to the driver		These features provide steering AND brake/acceleration support to the driver		These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met		This feature can drive the vehicle under all conditions			
Example Features						• automatic emergency braking • blind spot warning • lane departure warning		• lane centering OR • adaptive cruise control		• lane centering AND • adaptive cruise control at the same time		• traffic jam chauffeur		• local driverless taxi • pedals/steering wheel may or may not be installed		• same as level 4, but feature can drive everywhere in all conditions	

Figure 2-1 Society of Automated Engineers J3016 Levels of Driving Automation (SAE International, 2021)

At SAE level 3, automation capability extends to the monitoring task, allowing drivers to switch their attention towards NDRTs. However, the human driver remains responsible for the vehicle's actions and must be ready to intervene in the event of a system failure or boundary limitation (Large et al., 2019). The role of the driver will change depending on who has control of the driving task. Therefore, drivers will not only need to build up the requisite skill set to effectively interact with the automated system, but they will also require an additional set of skills to be able to smoothly transition between automated and manual driving modes (Gold et al., 2018).

A key challenge in system automation is the inverse relationship between automation and human performance (Banks and Stanton, 2016, 2019). For example, when decision making functions become automated, the driver naturally gives less attention to the driving task. This unintended consequence of automation (Parasuraman et al., 2000) takes the driver 'out of the loop' (OOTL) of control; reducing their level of perception and comprehension of the system state and driving environment and the projection of their future state, a construct termed 'situation awareness' (SA) (Endsley, 2017a). Empirical studies investigating transitions from automation to manual control (e.g. Dogan et al., 2017), have highlighted performance challenges associated with the re-engagement of drivers' cognitive and perceptual-motor controls necessary to effectively takeover the driving task, suggesting that when a system issues a takeover request (TOR), the driver may not be ready to drive. Therefore, to realise the full potential of current and future automated vehicles, drivers need to learn how to operate and interact with them in a safe and appropriate manner (Beggiato et al., 2015).

2.2 Lessons from Aviation

A common assumption, and often a key selling point, regarding the steady integration of automated systems, is that the more tasks that are delegated to automated control, the easier it will be for the human to operate the system and the less training they subsequently require (Cummings et al., 2019).

Within the aviation sector, a large body of human factors research (e.g. Bainbridge, 1982; Sarter and Woods, 1994; Parasuraman and Riley, 1997) demonstrates how the introduction of automation fundamentally changed the flying task and role of the pilot. Research findings highlight the misunderstanding and confusion that automation created for pilots attempting to operate new, complex human-automation system, and the deleterious effects this had on pilot performance and error (Casner and Hutchins, 2019). These findings were subsequently incorporated into the design and delivery of rigorous training programs to provide pilots with a more complete understanding of the effects of automated systems. Even so, reports regarding the recent incidents and fatal crashes involving the Boeing 737 Max (Campbell, 2019) suggest that a contributory factor was a lack of pilot knowledge and understanding of the automated system that had been incorporated into the aircraft. It is alleged that pilots were given minimal provision of computer-based,

self-administered training on the 737 Max, which did not include the additional automated system. Consequently, pilots were not equipped with an understanding of the rules or procedures to respond to the system when it failed. Although the pilots were able to use their considerable knowledge and experience to mitigate the impact of the first reported incident of a system fault, there followed two fatal crashes involving failure of the same automated system. This example suggests that, even in a sector where there is extensive research to the contrary, there remains a gap in appreciation for how new automation impacts human-automation interactions and shared control of the system and highlights the importance of updating the mental model of the human agent.

It is noted that in the current driving context, no such training provision is provided for publicly available level 2 vehicles (other than the prerequisite user's operating manual). In light of this and a recent announcement by the USA National Highway Traffic Safety Administration that a special investigation will be launched regarding vehicles with assistance systems following further fatal automation-related incidents involving these vehicles (Shepardson, 2020) (for example, see National Transport Safety Board, 2020; 2019), there would appear to be a need to review the training needs for drivers of automated vehicles with a similar vigour as that provided within the aviation field.

2.3 The changing driving task

The driving task is understood to involve many individual sub-tasks. Michon's (1985) hierarchical model of the driving task categorises sub-tasks into three behavioural levels: control (or operational), tactical (or manoeuvring) and strategic (Figure 2-2).

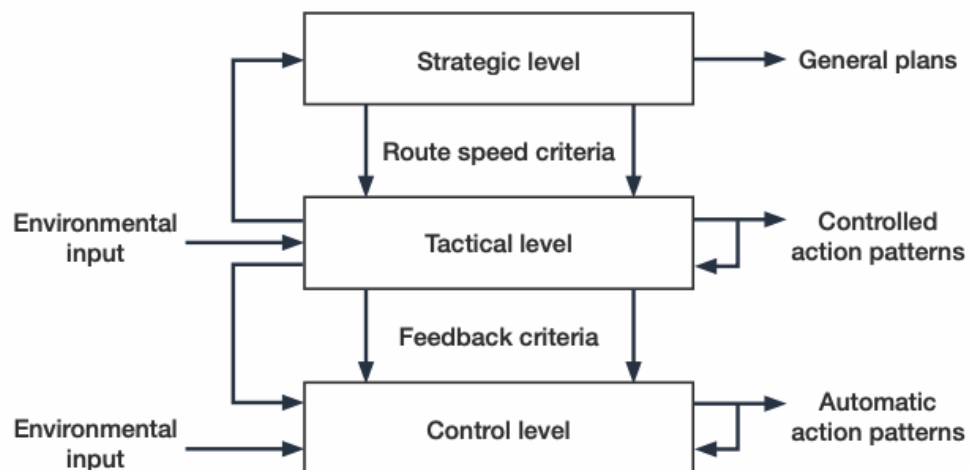


Figure 2-2: Michon's (1985) Driving skills hierarchy

At SAE level 2, automated systems are capable of intervening in control level tasks. Drivers of vehicles fitted with partial driving automation, for example, can potentially enjoy increased comfort and safety as their

cognitive resources are freed up through the delegation of menial, primary control actions to features such as active steering (AS) working alongside adaptive cruise control (ACC) (Kircher et al., 2014). As the sophistication of automation increases to level 3, system control capability will extend to tasks at the tactical level. In addition to level 2 capabilities that can control a vehicle's lane position, speed and a set timed headway to a lead vehicle, ACC systems will be able to monitor the driving environment, enabling rule-based decisions that allow safe interactions with other vehicles based on the interpretation of the immediate situation (Kircher et al., 2014). However, any vehicle with less than full automation capability still requires the human driver to remain in the control-feedback loop, playing an active role in the driving task (Banks and Stanton, 2019). For example, at level 2, the human driver is required to continuously monitor the system and driving task and be prepared to take over at any point. At level 3, although the human driver has been taken out of the loop of control by design (Merat et al., 2019), the ODD is bounded. Therefore, the driver remains actively responsible for the vehicle and is, consequently, required to be able to take over from the automated system should it be required due to system operational limitations or failures.

Using automation to control parts of the driving task fundamentally restructures the task as a whole, bringing with it changes to the role and responsibilities of the human driver (Banks et al., 2014; Kircher et al., 2014; Banks and Stanton, 2016). However, the new driver's role has not been formally and specifically outlined within the available automation taxonomies, leaving the skills, knowledge and potential training needs of the human driver in sharing the driving task with automated systems open to interpretation. The role of the driver within an automated vehicle and the challenges related to this changing role is considered to be analogous to that of a pilot in an automated aircraft (Stanton and Marsden, 1996; Casner and Hutchins, 2019).

Consequently, the roles of pilot-flying, and pilot-monitor, as applied in aviation (Casner and Hutchins, 2019), have been transposed to the driving domain at the intermediate stage of automation development, reflecting the changing, rather than diminishing, responsibility of the human driver as they share control of the driving task with automation (Banks and Stanton, 2016; Banks, et al., 2019).

The concept of 'shared control' is often used to reflect the cooperative requirements of the driving task in automated vehicles (Flemisch et al., 2012; Large et al., 2017a; Banks et al., 2019). This highlights the critical role that both the automation and the human driver play in the successful completion of the driving task. It draws particular attention to the importance, complexity and challenges of shared situational awareness within a joint cognitive system that is continuously evolving. The shared control approach to the design of automated vehicles relies on the driver's understanding of system capabilities and limitations for effective human-machine cooperation (Large et al., 2017a; Banks et al., 2019). Therefore, there is a requirement for both transparency and clarity of the shared driving role (Flemisch et al., 2012) and the

identification of training needs and certification (Kyriakidis et al., 2019) that will enable a dynamic balance to be maintained between the control, ability, authority, and responsibility of the actors within the system (Flemisch et al., 2012). Until automation capability requires no human input (i.e. SAE level 5 cars, or SAE level 4 pod-type vehicles in geo-fenced ODDs), the human driver remains a critical agent within the driving task (Banks et al., 2019). It is therefore important to investigate not only how the role of the human driver changes with the introduction of differing levels of automation, but also the dynamic nature of the human driver's role and responsibilities during any one journey, so that necessary training or guidance can be provided through appropriate means (Banks and Stanton, 2016).

2.4 Visual attention and situational monitoring

Driving is primarily a visual task (Merat et al., 2019) and deficiencies in visual attention are widely reported to be responsible for a large proportion of road traffic accidents (e.g. Klauer et al., 2006). Driver inattention can be considered the mismatch between how the driver has allocated their attentional resources (and to what activities) and the resources demanded by activities critical to safe driving, inclusive of any sensory, perceptual, motor and cognitive mechanisms (Lee et al., 2009). For example, so called 'looked, but failed to see' accidents, are attributed to a mismatch between a driver's allocation of attention to a road scene, and their selective (in)attention to a particular feature, such as a cyclist, within their visual array – a phenomenon known as 'inattention blindness' (Herslund and Jørgensen, 2003). A key task for drivers at all levels of the driving hierarchy is that of situation monitoring - continually processing perceptual information to generate an understanding and prediction of dynamic changes in the driving environment for use in physical vehicle control. It is this coupling of the physical control of the vehicle and the situational monitoring (situation awareness) that has recently been used to clearly define the control loops as mapped onto Michon's (1985) hierarchical model of the driving task, including the different spatiotemporal scales associated with control of subtasks at the operational, tactical and strategic levels (Figure 2.1). For example, continuous monitoring is required for vehicle operations, whereas monitoring relating to tasks at the tactical level varies in response to characteristics of the driving environment (Merat et al., 2019) (see Figure 2.2).

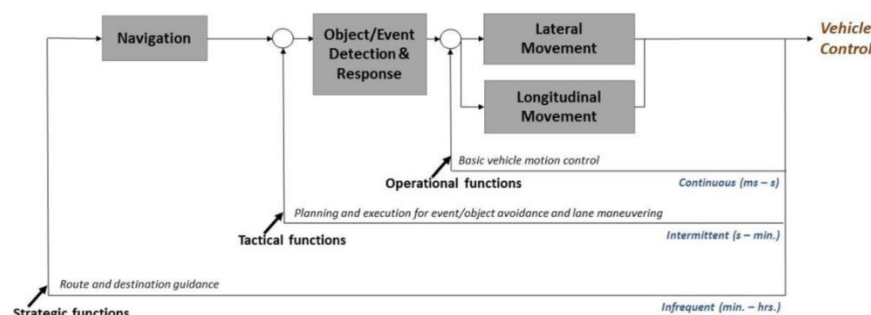


Figure 2-3 Multi-level control in driving, including situation monitoring at each level (Merat, 2019)

The associated 'in', 'on' and 'out of the loop' definitions, alongside this model, are useful when thinking about the interrelationship between the driving task, the role of the driver and level of vehicle automation. This is particularly relevant when identifying the impact of increases in automation capability on drivers' perceptual-motor control loop and the resultant implications for informing safe and appropriate skills and behaviours for drivers of these vehicles and required guidelines for system use. According to these definitions, 'in the loop' is in physical control of the vehicle and monitoring the vehicle, 'out of the loop' (OOTL) is defined as not in physical control of the vehicle and not monitoring the driving situation OR in physical control of the vehicle, but not monitoring the driving situation, and 'on the loop' is not in physical control of the vehicle, but monitoring the driving situation.

A key requirement for drivers of level 2 vehicles is to continually monitor the situation while the vehicle systems are in physical control of the vehicle. Consequently, this level of automation demands the driver remains 'on the loop' during automated periods of driving. However, as is well documented within the literature, a predominant consequence of level 2 automation lies in drivers entering an OOTL state due to challenges associated with this new driving task of passive monitoring. Mitigating strategies should therefore be geared towards supporting drivers in remaining 'on the loop' when using the automated systems. Additionally, vehicle designers need to ensure that drivers are given adequate and correct guidelines to gain accurate understanding of the functionality, capability and limitations of the automation, including the consequences of being OOTL, so that they know how to safely use the system (Carsten and Martens, 2019; Merat et al., 2019). In contrast, the changes in the driving task and driver role at level 3 automation induces OOTL driver state during automated driving, by design. Consequently, a key part of the driver role at this level is the smooth transition in and out of the loops of control in accordance with automated and non-automated driving modes. The reduction in SA and attention at this level of automation is not necessarily unsafe (Carsten and Martens, 2019), as long as the driver is able to calibrate their levels of SA and attention to accurate system reliability. Therefore, in order to ensure safe use of these systems, drivers need to understand what level of SA and attention they should have in relation to different modes of automation and how they need to interact with the system to adjust and tune their SA and attention in a timely manner during dynamic operations.

2.5 Information processing and cognition

Rasmussen's (Rasmussen, 1983) skills-, rules- and knowledge-based behaviour model (SRK) presents a hierarchical taxonomy, similar to Michon's (1985) model of the driving task (Sullivan et al., 2016), which links information processing and cognition to increasing levels of complexity (Cummings, 2014). At the foundational level, skill-based behaviours are sensorimotor actions, such as those required for control level driving tasks. Once acquired and mastered with sufficient training,

skills become highly automatic, freeing up attentional resources for higher cognitive tasks. At the intermediate level, rule-based behaviours are actions guided by 'if-then' rules, accumulated from previous training and experience and formed through associations between cues and the appropriate action selection (Wickens et al., 2014). Rules are stored either as mental structures, termed 'schemas', or as written procedures. The operator selects the rule by processing input against system state. For example, at the tactical level of the driving task, decisions to give way to or overtake another road user is based on the driver's interpretation of the situation using input monitoring and SA of the traffic environment and automated systems. Issues arise when the human operator does not recognise the correct goal and selects the incorrect procedure. When facing novel situations, operators do not have existing rules to guide action selection and need to work at the knowledge-based level of behaviour. Mental models are built as the operator gains knowledge and experience of the system, creating a set of expectations and conceptual information about the system that can be used to guide the selection and formulation of an action plan for an explicit goal (Cummings, 2014).

For experienced drivers, the actions required to control the moment-to-moment operations of a manual vehicle are predominantly skill-based behaviours – automatic (Kircher et al., 2014; Merat et al., 2019), well-practised, and performed without effort or conscious thought. However, in the context of the automated driving task, automatic performance of these behaviours, based on the mental model of manual driving, may no longer be appropriate. In other words, as with the pilots of Boeing's 737 Max (Campbell, 2019), drivers of these vehicles need to understand the new 'if-then rules' introduced to the driving task through automation, and the perceptual-motor actions or 'skills' they need to apply in the context of automated driving.

Arguably, there is a need to understand how drivers apply skills learnt within the manual driving context to the performance of new tasks required in the automated driving context, such as transitions of control. Exploring how drivers interact with these future vehicles could help to reveal potential issues and inform requirements in relation to driver training or support from interface design to encourage the uptake or maintenance of desired driver behaviour for the automated driving task. Recent work summarising potential consequences of automation linked to OOTL state from the aviation and automation domain (Seppelt and Victor, 2016) include items such as: passive monitoring, failure to some safety critical areas, e.g. rear-view and side mirrors; high trust scores; low SA scores; and inaccurate mental models. As well as providing themes to address with training or interface countermeasures, these items also present options for quantitative and qualitative measures by which to judge driver performance relating to the induction of OOTL driver state during automated driving and the effectiveness of countermeasure strategies.

2.6 The changing needs of the human driver

Vehicle automation changes the knowledge and understanding the human driver needs in relation to the driving task and operational working of the vehicle (Casner and Hutchins, 2019). Previous technological developments provided general performance improvements to vehicle functions that were already available and managed by the human driver (Sullivan et al., 2016). For example, automatic cruise control (ACC) holds a vehicle's speed as selected by the human driver. This automatic function removes the need for constant driver input to the accelerator pedal, thereby providing the driver with improved control and comfort. However, these improvements do not significantly change the driving task or the role for the driver, who remains in control of operational tasks, such as acceleration and deceleration, as well as managing the vehicle's interaction with objects in the external environment, such as pedestrians and other vehicles, through steering, lane selection, etc. Moreover, any input from the driver to the pedals typically disables the cruise control function.

Automation capability provides the human driver with an array of tools to which they can delegate parts of the driving task. Advanced Driving Assistance System (ADAS) features such as, blind spot detection and lane departure warnings play an active role in the overall driving task, controlling tasks at the operational level of driving and supplementing tasks at the tactical level (Casner and Hutchins, 2019). With previous vehicle developments, responsibility for the overall management and accountability of the driving task remained, unambiguously, with the driver. However, in the new human-machine context, the distribution of responsibility requires clarification and reiteration. Within their role, drivers command and control the use of automated features within the driving task. For this to be successful, they need to understand, and subsequently manage, how, when, and why they use them. This level of proactive thinking sets apart the cognitive requirements for drivers of automated vehicles and is facilitated by the appropriate knowledge and understanding of the automated driving task (Casner and Hutchins, 2019). Casner and Hutchins (2019) propose that the knowledge required of drivers of automated vehicles is akin to that of pilots using cockpit automation. They refer to Degani and Weiner's (1994) 3Ps of flight deck operations, suggesting that, as with pilots, drivers require:

- *Automation Philosophies* – the need to understand their reasons for using automation;
- *Automation Policies* – the context dependency of when automation should be used; and
- *Standard Operating Procedures* – outline how automation should be used.

Together, these three concepts indicate that there is an additional depth of understanding required by the driver to enable the interaction with automated features and get maximum benefit from their design, above

and beyond technical operation. However, there is still a requirement to translate this into the skills, rules and knowledge required by drivers.

2.7 Behavioural adaptations and procedural deviations

Parasuraman et al., (2000) state that automation does not replace human activity, but rather changes it, often in accidental and surprising ways. Research into behavioural adaptations of human drivers interacting with automated vehicles have highlighted a number of issues relating to aspects of performance and safety, similar to those seen when automation was introduced in the aviation domain (Bainbridge, 1982).

Naturalistic studies examining driver experience with level 2 automated vehicles have reproduced evidence of phenomena such as ‘automation bias’, a concept that described pilots’ reduction of risk perception, alongside over-reliance on automated warning systems, due to a lack of understanding regarding the capability and competence of the automated systems being used e.g. (Brown and Laurier, 2017; Endsley, 2017; Banks et al., 2018). Qualitative analysis of driver interaction with these cars has provided anecdotal evidence of increased engagement with non-driving related tasks (NDRTs), as well as reduced situation awareness and vigilance, impacting upon the ability of the human driver to maintain shared control with the automated system. For example, video analysis of Tesla’s ‘Autopilot’ showed multiple drivers testing the limits of the system and driving for extended periods with no hands on the steering wheel, despite explicit knowledge that the assisted driving provided is not ‘hands-free’ (Banks et al., 2018). In one instance, this lowered risk perception was coupled with ‘mode confusion’, where a driver failed to notice they were driving hands-free in manual mode. Evidence of drivers’ reduced vigilance of the traffic environment and the state of the automated system was shown in drivers failing to notice and correct ‘Autopilot’ errors, such as a misrecognition of exit lane lines. In this example, the autopilot’s misinterpretation was visible to the driver on the dashboard, but the driver did not notice until the vehicle had taken the wrong exit. Interviews carried out with Tesla drivers by Lin et al. (2018), revealed driver understanding of the system’s capability (their ‘mental model’) was often incorrect, resulting, for example, in over reliance on the system and encouraging them to engage with their NDRT rather than monitor the roadway. Although this reported patterned use of trial-and-error demonstrates learning by understanding the car through its behaviour, this has been shown to encourage drivers to deviate from the desired operating procedures of these systems and has important road safety consequences.

At SAE level 3, the extension of automation capability to the monitoring task allows drivers to engage with NDRTs until the operational design domain for the system is reached, or the system fails and the driver needs to re-engage with the driving task. This increase in automation capability changes the way the human-machine system interacts,

further impacting the role of the human driver. For example, at level 3 automation, the perspective on driver engagement with NDRTs shifts from one of driver distraction, to task interruption and task switching (Janssen et al., 2019). Transitions of control of the driving task between the automation and driver during dynamic operations introduces an additional set of challenges in relation to driver behavioural adaptation, increasing the potential for procedural deviations that could lead to OOTL issues that negatively impact the safety and effectiveness of the transitions following system takeover requests (TORs).

Research into controlled transitions from automation to manual driving have looked to determine the optimal transition time for drivers to safely and comfortably take over from the driving task, and have proposed between 3-10 seconds, typically based on the mean or median response times observed (Melcher et al., 2015; Merat et al., 2014). More recently, Eriksson and Stanton (2017) highlight the need for adaptive automation – professing that TOR lead times will depend on context and, consequently, suggesting that transition times could extend to over 25 seconds in some situations. Even so, for reasons of experimental control and internal validity, research studies tend to use a pre-defined NDRT that users engage with throughout the period of automation. Although this provides standardisation from which to compare takeover performance across participants, this experimental rigour potentially omits important motivational factors that could have a significant bearing on driver behavioural adaptations. For example, when humans divide their attention between two competing tasks, intervening factors such as prioritisation and motivation will determine time-on-task. During periods of automation control, the human driver may subsequently prioritise their NDRT over interaction with the driving task, regardless of their fall-back responsibility, delaying re-engagement with situation monitoring required to get back in the loop of control.

A longitudinal study conducted by the University of Nottingham in collaboration with the RAC Foundation (Burnett et al., 2019; Large et al., 2019), was designed to investigate the types of activities naturally undertaken by drivers during periods of automation in SAE level 3 vehicles. The study invited participants to undertake five simulated journeys over consecutive days. The journeys included transitions between manual and automated driving and an extended period in 'automation mode'. Participants were given free choice about the activities they undertook during automation mode. As the choice of NDRT was unrestricted and uninfluenced, it was possible to investigate any mediating effects of personal motivation on task prioritisation and task switching, and analyse any resulting behavioural adaptations of human drivers, in particular, during the transition from automation to manual driving mode. Findings showed that drivers placed greater emphasis on learning how to effectively re-engage with situation monitoring and physical control at the operational level, whilst neglecting to engage with situation monitoring at the tactical level until after physical control of the vehicle had been transferred, despite explicit instruction to 'prepare to drive' as part of the system takeover

request (TOR) and exposure to a failure scenario (poor visibility due to fog). Instead, as drivers gained experience and knowledge of the system, they gave priority to their NDRTs. This suggests that, at the point of manual mode engagement, drivers may not be 'ready' to carry out tactical level tasks safely and effectively and resume complete control of the driving task. It highlights a need to design training or system design strategies that not only increase driver knowledge of system capability and limitation, but also clearly defines essential operating procedures that guide required driver behaviours, thereby facilitating the safe and effective flow in- and out- of the loop of control and mitigate against procedural deviations that could lead to road traffic accidents or near misses relating to driver inattention.

Research concerning behavioural adaptations echoes previous findings within the aviation domain, where pilot operational deviations were attributed as the leading cause of aviation crashes (Degani and Wiener, 1997). They emphasise that operating procedures form an integral part of the activities required to ensure successful operations in complex human-machine systems. These procedures provide a way to standardise and specify the manner in which the required tasks should be carried out, giving clear instruction to the human operator to ensure tasks are carried out in an optimal, logical, safe and predictable way. A well-designed operating procedure should optimise the sequencing of tasks and promote efficient scheduling by the human operator (Degani and Wiener, 1997). In the context of transitions of control at level 3 automation, this procedure should arguably include sequencing tasks associated with getting the human driver 'on-the-loop' of control at the tactical and strategic task levels, before the driver regains physical control of the vehicle. Alongside operating procedures that provide adequate and correct guidance to drivers on how they should be using the system, research findings on behavioural adaptations suggest there is also a need to inform drivers about the effects of automation on their own behaviour and the potential consequences of being OOTL to support safe interactions between the human driver and the automated systems (Carsten and Martens, 2019; Casner and Hutchins, 2019; Merat et al., 2019). The development of strategies focused on supporting the uptake of desired behaviours and mitigating undesirable behavioural deviations, associated with reduced SA and vigilance and increased NDRT use, should consider both user interface design and training interventions to investigate countermeasure solutions that are necessary, sufficient and versatile in order to meet the wide array of requirements from the multiple stakeholders within the automotive field. In particular, interface design solutions are arguably appealing given the costs and complexities involved in rolling out remedial training for experienced drivers. Given that it is not the skills necessarily, but the procedural rules involved with situation monitoring in the automation context, that are novel to this group of drivers, it is suggested that use of an operating procedure that can be built into the design of an HMI would provide a desirable solution.

2.8 Mental models, expectations and trust

Vehicle automation offers novel and complex technology. Research findings highlighting the ways humans interact with these automated vehicles demonstrate the importance of the type, level, and quality of training and knowledge imparted to drivers, who do not yet have an accurate mental model of system functionality and limitations (Beggiato and Krems, 2013; Beggiato et al., 2015; Casner and Hutchins, 2019). Mental models are formed during the learning process and create a set of expectations in relation to what the system contains, how it works and why it works that way (Beggiato and Krems, 2013). Research has shown that incorrect mental models lead drivers to make overly optimistic assumptions about automation capability, influencing how they interact with and manage the system. This, in turn, can lead to 'automation surprises', such as unexpected system responses or failures, which can impact trust and acceptance. For example, a study by Beggiato and Krems (2013) found that the provision of different descriptions of an adaptive cruise control (ACC) system influenced driver mental models and expectations of system functionality. Following simulated drives using ACC, the mental models of all groups converged towards realistic system functionality. However, trust and acceptance scores for the incomplete group showed a steady decline following driver experience of unexpected emergency scenarios. These results highlight the importance of clear and consistent learning strategies to the development of accurate mental models from the outset.

Trust is an important factor in influencing how safely and effectively drivers interact with complex automated vehicle systems (Parasuraman and Riley, 1997). The presence or absence of trust is not dichotomous, it is a dynamic phenomenon that increases or decreases based on the driver's perceptions of how the vehicle system operates and whether their beliefs attached to these perceptions give rise to positive or negative attitudes (Walker et al., 2016). Limited and inconsistent knowledge provided to human drivers about the underlying principles and mechanisms behind automated functions creates uncertainty and variability in the perceptions, expectations, beliefs, and, therefore, trust in the automated system. A lack of formalised training for automated vehicles means that drivers are provided with limited and incomplete knowledge of automated functions (Casner and Hutchins, 2019). The lack of clarity and consistency in drivers' understanding of underlying principles and mechanisms of automated systems creates uncertainty and variability in their mental models and resulting trust, which influences how the driver behaves with the automation (Beggiato and Krems, 2013; Walker et al., 2016). If a driver's trust exceeds the system's capabilities, they are likely to over-trust, and over-rely on the automated systems, something Parasuraman and Riley, (1997) termed 'automation misuse'. Over-trust results in both risk perception and attention levels being reduced, making drivers susceptible to monitoring failures, slow reaction times and poor reaction quality to critical events (Körber et al., 2018).

Empirical evidence from level 3 driving simulator studies has strengthened the arguments for the provision of formalised training that fosters accurate mental models in drivers (Nylen et al., 2019). For example, Korber et al. (2018) demonstrated the influence of introductory information of system capabilities on trust and behavioural outcomes. Participants provided with 'trust-promoted' information, prior to completing critical and non-critical simulated takeovers, showed a higher number of collisions, deteriorated takeover performance and increased engagement with NDRTs in comparison with the trust-lowered group. These results show the safety implications of creating a sense of 'over-trust' with automated systems and the potential safety benefits of developing appropriately calibrated trust. The authors posit that early training interventions that emphasise how interactions with the dynamic driving environment influence the capability of the system to cope with a situation are required to mitigate the negative effects of automation failure. This idea supports the recommendations from Walker et al. (2016), which finds that to establish trust in vehicle automation that promotes acceptance and successful implementation of the technology, drivers need to be made explicitly aware of what the system is designed to achieve.

These findings highlight how training processes that impart clear and accurate information about the characteristics and capabilities of the automated system from the outset are essential to fully harness the potential safety benefits of this level of automation (Beggiato et al., 2015; Sullivan et al., 2016). However, current training methods for consumers of level 2 autonomous vehicles is often self-directive, passive and inconsistent. According to a recent survey (McDonald et al., 2018), consumer training tends to happen via either: user manuals, where learning effectiveness depends on the user reading, understanding and remembering operational procedures (Casner and Hutchins, 2019); or instructions from sales personnel at car dealerships (Abraham et al., 2017), which have been reported to vary widely in quality depending on personnel training and capability to explain technology to prospective buyers.

2.9 Driver Competence, Capability and Managing Driving Task Demands

At an early stage of learning to drive, skill-based tasks such as operational vehicular control (Michon, 1985) use controlled attentional processes (Shiffrin and Schneider, 1977). Proficiency of these operational tasks provides capacity for the driver to focus on building competence in rule-based behaviour. This includes tasks such as visual search and pattern identification of road user behaviour, which makes up the predominant behaviour in driving (Fuller, 2000). Rule-based behaviour is driven by a set of rules or conditional statements (Rasmussen, 1983; Cumming, 2014). Drivers learn these rules either directly, via documentation such as the highway code, or, more likely, through contingent learning by experience on the road, i.e. learning through consequences. However, in an event where there is no rule to

guide behaviour, drivers need to refer to their broader mental representation to formulate their decision making and response action (Fuller, 2000; Rasmussen, 1983; Cummings, 2014; Teh et al. 2014).

Level 3 automated driving system (L3) (ADS) capability introduces the requirement for planned control transitions (PCTs). These are requested of the driver due to the system approaching the pre-set boundaries of its operational design domain¹ (ODD). PCTs constitute a new subtask within the driving task and, as yet, an explicit set of behavioural rules to carry out PCTs does not exist. This is, in part, because it is assumed experienced drivers are already equipped with the competencies to handle these transitions, either through rule-based behaviour already learned or the application of knowledge-based behaviour from driver mental models. As evidenced in the thesis so far, drivers are capable of rapidly developing strategies to quickly develop proficiency in taking over control of steering and speed control with the required level of skills-based competence during control transitions, both planned and failure-induced (Large et al., 2019; Shaw et al, 2020a). However, without explicit rules and guidance on the behaviours required to regain control of the tactical level tasks involved in driving, PCTs are reliant on driver's knowledge-based behaviour to safely and effectively perform this task. The thesis proposes that currently drivers do not understand or consciously recognise the need to conduct the behaviours required to effectively build situational awareness (SA), while the ADS remains in control of the driving task. Additionally, drivers may also expect the in-vehicle HMI system to notify them of relevant aspects within the driving environment and rely upon these systems to direct their attention, rather than proactively gathering this information themselves. These issues, relating to skill-degradation, complacency and automation bias, are already well-known within the aviation domain (e.g. Brown, 2016). They have the potential to degrade the momentary and long-term driver competence of key skills required for proficient execution of tactical level tasks. In addition, reliance on the HMI provides no redundancy for automation failures both in terms of sensing and communication within the joint system.

2.9.1 Task Capability Interface model (Fuller, 1999)

Fuller (1999) defined driving as the management of task difficulty. According to Fuller's 'Task Capability Interface model', a driver's perception of task difficulty depends on the transactional relationship between the level of task demand by the external environment and the momentary capability of the driver. Importantly, as the driver's

¹ in contrast to failure-induced control transitions, which are requested due to a system failure or critical situation that cannot be dealt with by the ADS. Both types of control transition are possible during L3 automation. However, it is anticipated that at this level of system capability, a failure-induced control transition will be an unusual rather than typical event. Additionally, if the driver is unable to take control or unresponsive to the TOR it is expected that the ADS will perform a minimal risk manoeuvre.

perception of task difficulty increases so does the level of risk. This, in turn, can lead to a greater likelihood of driver error, loss of control, and the probability that an error will lead to a crash. Driving is a dynamic task and changes in task demands can be sudden, urgent and unpredictable (Teh et al, 2014). Fuller (1999) defined task demand as the objective complexity of the task, arising from a combination of features of the driving environment. For example, the behaviour of other road users; conditions of the driving environment (e.g. traffic density); the performance and control of the vehicle; and aspects such as the vehicle speed and position; and the communication of the driver's intended actions (See **Error! Reference source not found.**). The higher the task demands, the higher the level of momentary driver capability required to meet the threshold to adequately manage the driving task.

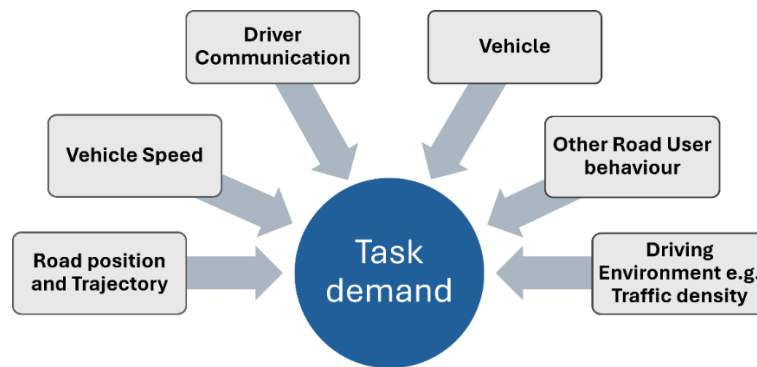


Figure 2-4 The features contributing to 'Task Demand', according to Fuller's (1999) Task Capability Interface model

On the other side of the task demand relationship is driver capability. According to Fuller (1999), driver capability is constrained by driver competence, which is determined by factors such as driver experience, training and education, as well as individual differences such as perceptual acuity or co-ordination. Driver competence covers a range of skills including vehicular control, hazard detection and recognition (i.e. the ability to 'read the road'), and anticipatory and defensive driving skills. A driver's level of competence develops from processes such as training, formal learning and experience (Fuller, 2000). It is also shaped by internal factors such as motivation (e.g. adherence to rules) and expectations influenced by driver mental models (Teh et al. 2014). However, driver competence is dependent on the contextual ability of the driver to deliver their level of competence at a given moment i.e. driver capability (Fuller, 2000, 2005). A key characteristic of driver capability is situational awareness (SA), the ability to correctly predict the immediate future state of the unfolding road and road user scenario ahead (Fuller, 2005; Endsley, 2005). Experienced drivers achieve this through behaviours and skills that allow anticipatory control responses (Fuller, 2005). For example, the use of timely and efficient allocation of attentional, decision making and control response, resources to effectively and proactively adapt to the dynamic task demands of the driving environment (Fuller, 2000; Endsley, 2017a; 2018). Higher levels

of SA are associated with anticipatory hazard avoidance through increased attentional focus, level of vigilance, speed and directional adjustments etc. being associated with higher levels of experience (Fuller, 2005), whereas drivers with low SA typically demonstrate more reactive behaviours when dealing with hazards (Endsley, 2018).

Driver capability is limited by competence and shaped by variations in driver state, such as, fatigue, distraction and mental workload (i.e. the amount of information processing capacity used for task performance etc.) (De Waard et al., 2008; Brookhuis et al. 1991; Fuller, 2005). At L3 automation, the intentional design, to allow drivers to turn their attention away from the driving task, means that drivers' well learned performance patterns are disrupted (Endsley and Kiris, 1995). The resultant reduction in cognitive and motor engagement with the driving task lowers SA and workload, inducing a passive, rather than active driver processing state (Petermeijer et al., 2015; Endsley, 2017a; Endsley and Kiris, 1995). Driver capability in this context will, therefore, be shaped by whether the driver is in-, on-, or out- of the loop of control following a request to transition to manual driving control (Merat et al., 2019) and how calibrated their level of SA and attention is in relation to the task demands of the driving environment (Endsley and Kiris, 1995; Endsley, 2018; Carstens and Marten, 2019). Driver capability is also mediated, positively or negatively, by human factors, such as situational awareness and mental workload, and internal factors, such as motivation or adherence to rules. These factors together will determine the momentary capability available to the driver to meet the variable demands of the task (**Error! Reference source not found.**).

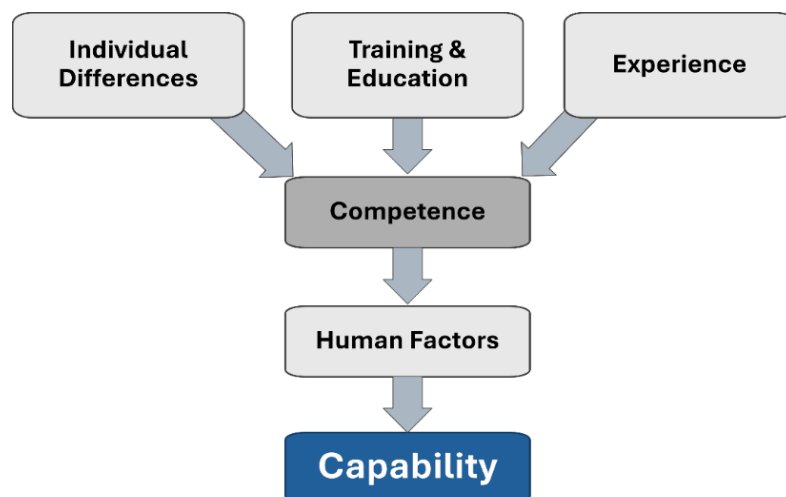


Figure 2-5 The Factors contributing to driver 'Capability', according to Fuller's (1999) Task Capability Interface Model

A mismatch in the external task demands and the available driver capability impacts the driver's perception of task difficulty and the level of risk, with both increasing as the threshold of the driver being unable to meet the demands of the task is approached. For example, if the task demand (D) is high and the momentary driver capability (C) available is

low ($D > C$) the task will be perceived as (too) difficult. However, if the available driver capability is greater than or equal to the task demands, the task will be perceived as manageable or even easy ($C \geq D$) (see **Error! Reference source not found.**).

Figure 2-6 The transactional relationship between Task demand and momentary driver capability (Adapted from Fuller, 2000).

The Fuller (1999) Task Capability Interface model is a useful tool to highlight the potential challenges of introducing L3 automation with no additional training or guidance for experienced drivers. Firstly, if, as this thesis argues, PCTs constitute a new sub-task for the driving task in the context of vehicles with L3 automation, then a lack of training or guidance on this sub-task will constrain driver competence. Secondly, if the calibration of human factors such as SA and attention during the planned control transition period (PCTP) is inappropriate, this alone could negatively impact the available capability to the driver from the point manual driving is resumed. As outlined by Fuller's (1999) model, this reduction in capability may not be an issue if the available capability is greater than or equal to the task demand (albeit this, in itself, may introduce issues of complacency). However, should the task demand be (or become) high and the momentary driver capability remain low then the driver perception of task difficulty (and, consequently, driver mental workload) will also be high as will the risk of driver error and likelihood of a crash or the need for evasive action.

Driving is a highly dynamic task involving vast arrays of information (Teh et al, 2014; Brookhuis et al, 1991; Fuller, 2000). The task demands of driving depend on the dynamic combination of environmental features such as traffic complexity, other road users' behaviour, characteristics of the vehicle, and its speed and position on the road (Brookhuis et al. 1991; Fuller, 2000; Teh et al. 2014). For

example, a high task demand scenario might include features such as a curvy road, wet weather conditions, reduced visibility, in combination with high traffic density and a critical scenario such as a lead vehicle braking suddenly. In contrast, a low task demand scenario might involve low traffic density, on a straight motorway, in clear, dry weather conditions. To safely execute the driving task in response to fluctuating task demands, drivers need to be able to select, perceive and interpret relevant objects and elements from the traffic situation. This will allow them to carry out effective decision making and action responses within the driving environment (Teh et al. 2014). The (worrying) assumption with L3 automation is that experienced drivers will be able to seamlessly jump back into performing these cognitive capabilities at a level appropriate for the task demands (Mole et al., 2019). However, drivers' disengagement from the sensorimotor control loop during automated driving, means that it is inherently difficult for drivers to understand when and how attentional resources should be allocated to adequately rebuild reduced SA and calibrate the perceptual-motor control loops (Mole et al., 2019; Endsley, 2017a; Carsten and Martens, 2019).

Drivers, regardless of experience, lack the implicit or contextual understanding to proactively rebuild levels of SA and attention in a timely manner following a TOR, ahead of resuming manual driving (Carsten and Martens, 2019; Mole et al., 2019). Imparting this understanding is a critical factor in the development of driver performance and control of these vehicles, especially as traffic demands can be sudden, urgent and unpredictable (Teh et al, 2014). Insufficient or inappropriate behaviour may result in the driver having a level of SA that is too low and workload that is too high at the point they resume manual driving. Consequently, drivers may lack essential perceptual information and capability to execute the required immediate or anticipatory control responses for the task demands (Fuller, 2000; 2005; Mole et al., 2019). Additionally, large, spikey variations in workload can result in deleterious effects on attentional processes, leading to sampling errors such as cognitive tunnelling (Endsley, 2005) or inattentive blindness (Herslund and Jørgensen, 2003). This mismatch between driver capability and task demands may be reflected in task performance and result in an increased potential for errors and poor safety outcomes (Fuller, 2005).

2.10 Summary of literature review

The literature identifies clear concerns associated with the introduction of vehicles with driving automation capabilities onto our public roads without first considering the needs and capabilities of the human driver. Empirical evidence shows that as decision making and control functions are increasingly automated; the driver naturally gives less attention to the driving task. Subsequently, there is an inverse relationship between automation and human performance. However, until vehicles are fully automated, and require no human intervention, the driving task is one

that is shared between driver and vehicle. Therefore, to realise the full potential of these future vehicles, *drivers must learn how to operate and interact with automated vehicles in a safe and appropriate manner*. The key points that emerge from the literature (and guide subsequent activities) are:

- The driving task is one in which the responsibility of the human driver is changing rather than diminishing as they are required to share control of the driving task with the automation.
- Drivers therefore require an additional depth of understanding to enable interaction with automated features and get maximum benefit from their design, above and beyond technical operation.
- Automation does not replace human activity, but rather changes it, often in accidental and surprising ways. This is typically due to a lack of understanding regarding the capability and competence of the automated systems being used, and can result in drivers' over-reliance on the system, a reduction in their situation awareness and vigilance, and encourage engagement in non-driving related tasks (NDRTs).
- At higher levels of automation, system capability allows drivers to switch their attention to NDRTs, inducing an out-of-the-loop (OOTL) by design. Therefore, it is essential that drivers learn how to interact with these vehicles safely and effectively to be able to smoothly transition in and out of the loop of the driving task.
- Empirical evidence from level 3 driving simulator studies has strengthened the arguments for the provision of formalised training that fosters accurate mental models in drivers, yet no training provision is currently provided to drivers of current, publicly-available level 2 vehicles (other than the prerequisite user's operating manual). The lack of formalised training means that drivers may possess limited and incomplete knowledge of automated functions and inadequate or incorrect understandings of how and when they should use the system. This can result in drivers relying on the technology in situations for which it was not intended, or potentially switching it off completely, thereby failing to realise any benefit.
- The rising number of automation-related incidents involving vehicles with level 2 functionality suggests that there is a need to review the training needs of drivers of automated vehicles with a

similar vigour as that provided within the aviation field, particularly as the level of automation increases.

- Planned Control Transitions constitute a new sub-task for the driving task in the context of vehicles with L3 automation. They require drivers to be able to seamlessly switch their attention from NDRTs back to the driving task as and when the needs of the ADS arise. Insufficient or inappropriate behaviour following a TOR may result in the driver having a level of SA that is too low and workload that is too high at the point they resume manual driving to cope with task demands. Drivers, regardless of experience, lack the implicit or contextual understanding to proactively rebuild levels of SA and attention in a timely manner following a TOR, ahead of resuming manual driving (Carsten and Martens, 2019; Mole et al., 2019). Imparting this understanding is a critical factor in the development of driver performance and control of these vehicles.
- The Fuller (1999) Task Capability Interface model is a useful tool to highlight the potential challenges of introducing L3 automation with no additional training or guidance for experienced drivers. The model highlights the need to balance driver momentary capability with the task demands of the road, which can be sudden, urgent and unpredictable. Driver capability is constrained by driver competence, which, in turn is influenced by training and guidance and human factors elements such as situational awareness. A reduction in driver capability in itself may not be an issue. However, if the task demand is or becomes higher and the momentary driver capability the risk of driver error and likelihood of a crash or the need for evasive action increases.

3. Analysing Driver behaviour with Level 3 Automated Driving Systems

3.1 Introduction

This chapter outlines the observation study carried out on data produced by a longitudinal study conducted by the University of Nottingham in conjunction with the RAC Foundation (Burnett et al., 2019; Large et al., 2019).

A core argument of the thesis is that tacit assumptions that experienced drivers require no additional training or guidance to support the transition to driving vehicles with higher levels of driving automation are misplaced. A key aim laid out for this research is to investigate driver interactions and behavioural adaptations to understand what can be learnt from how drivers interact with these future vehicles. The aim of this phase of the research is to analyse driver behaviour in order to, firstly, validate the concerns raised about the lack of active training provision to this population of drivers; and secondly, to use the findings to guide potential design solutions.

3.2 Background

Progression towards fully autonomous driving is steady. However, until system boundaries no longer exist, the driving task is one which is shared between humans and technology (Brown and Laurier, 2017). The Society of Automotive Engineers (SAE, 2016) categorises vehicle automation into 6 levels of ascending capability. Levels are classified by the extent of system intervention in vehicle control of the driving task and the requirement of the human driver to monitor system performance and resume control of the driving task where necessary (White et al, 2019). Michon's (1985) hierarchical driving model categorises the driving tasks into three behavioural levels: control, tactical and strategic. At SAE level 2, automated systems are capable of taking over control level tasks, for example, lateral and longitudinal vehicle controls. However, the human driver remains responsible for monitoring the system and must be ready to take over the driving task if required. At level 3, automation capability extends to tactical level tasks such as monitoring, allowing drivers to switch their attention towards non-driving related tasks (NDRTs). However, the human driver remains responsible for the vehicle's actions and must be ready to intervene in the event of a system failure or boundary limitation (White et al. 2019).

A key challenge in system automation is the inverse relationship between automation and human performance. For example, as decision making functions are automated, the driver gives less attention to the driving task. This unintended consequence of automation takes the driver 'out of the loop' of control; reducing the driver's level of

perception and comprehension of the system state and driving environment and the projection of their future state, a construct termed 'situational awareness' (Endsley, 2017b). This can be seen in the behavioural adaptations evidenced in recent naturalistic studies analysing the experiences of drivers operating the Tesla Model S, level 2 vehicle (e.g. Banks et al., 2018; Brown and Laurier, 2019; Endsley, 2017b). Qualitative analysis provided anecdotal evidence of increased engagement with NDRTs, as well as reduced situational awareness and vigilance, impacting the ability of the human driver to maintain shared control with the automated system.

At level 3, the extension of automation capability to the monitoring task allows drivers to engage with NDRTs until the operational design domain for the system is reached or the system fails, at which point the system will issue a takeover request (TOR) to the driver to re-engage with the driving task. Empirical studies investigating automation to manual control transitions have highlighted performance challenges impacting the ability of human drivers to effectively take over the driving task. Research has focused on determining the optimal TOR timing for drivers to safely and comfortably take over from the automation (e.g. Melcher et al., 2015). However, for reasons of experimental control, these studies tend to use a standard NDRT that drivers engage with during the simulated automation mode. Although this provides standardisation from which to compare driver take-over performance, this experimental rigour potentially omits important motivational factors that could have a significant bearing on driver behavioural adaptations at this level of automation. For example, the increase in automation capability changes the perspective on driver engagement with NDRTs, from one of driver distraction to task interruption and task switching (Janssen et al., 2019). When humans divide their attention between two competing tasks, intervening factors such as prioritisation and motivation will determine the time-on-task. During periods of automation control, the human driver may prioritise the NDRT over interaction with the driving task, regardless of their fall-back responsibility, taking them further out-of-the loop of control and impacting their readiness to take over control of the vehicle. Additionally, by allowing drivers to turn their attention towards NDRTs it interrupts the well-learned attentional patterns associated with maintaining situational awareness (Endsley and Kiris, 1995). For example, the duration and location (area of interest) of driver glances gives an indication of the visual behaviour they are using to build and maintain an accurate picture of the driving environment. Higher levels of situational awareness are associated with increased attentional focus and vigilance to the external road environment (Fuller, 2005). Accurate calibration of attention to the requirements of the driving task is important to be able to form appropriate expectations about future

events, including when it is safe to glance away from the roadway towards, for example, the dashboard (Seppelt et al., 2017).

Current concepts such as ‘complacency’ help to identify patterns in driver behavioural adaptation due to automation, which can influence the design and development of mitigation solutions, such as, alert systems or attention monitoring devices. However, concerns have been raised about the specificity of models to accurately predict human behaviour. For example, ‘complacency’ describes an overly optimistic assumption of satisfactory system states due to limited knowledge of safe and efficient modes of operation. This lowered expectation of failure is more commonly, but not exclusively, seen in novice users of automation (Kaber, 2018). This reflects the limited knowledge and experienced uncertainty that affords higher levels of cognitive decision-making behaviours (Cummings, 2018). However, an increase in knowledge and experience does not prevent human operational deviations that could be construed as complacency (Kaber, 2018), therefore, mitigating solutions targeting the current construct of complacency may be ineffective. Satisficing, known as the ‘good enough’ heuristic, is the behavioural tendency of humans to use the most accessible solution rather than the optimal one. This bias assists decision making and reduces strain on mental resources, increasing efficiency in operations. For example, driver reliance on automated blind-spot detection systems without the additional vigilance or monitoring behaviours required from the human driver (Sullivan et al., 2016). Use of this cognitive heuristic is often successful. However, in human-machine systems, where control is shared and dynamic in nature, the use of satisficing requires a level of error acceptance that reduces performance reliability (Kaber, 2018). It is argued that consideration of cognitive strategies and biases can increase understanding and precision of constructs which outline implications of automation on human behaviour. This will allow better identification and selection of mitigation strategies within systems design or driver training (Dekker and Woods, 2002).

3.3 Aims and Overview

The current study carried out analysis on an extensive existing data set from a recent longitudinal study (White et al., 2019; Burnett et al., 2019), where participants had carried out a series of simulated drives in a vehicle equipped with level 3 automated driving systems (L3 ADS). The aim was to investigate and interpret behavioural adaptations of human drivers during planned and failure-induced control transitions from automated to manual driving, carried out during simulated drives. Video data from the study, which investigated driver interactions in a vehicle with L3 ADS, provided a rich and immediately accessible data source. The study’s longitudinal design and driver-led choice of NDRT presented an opportunity to investigate how driver behaviour changed

over time, and any mediating effects of motivational factors on driver behavioural adaptations.

3.4 Method

Forty-nine experienced drivers took part in the original study (for full details see Large et al., 2019; Burnett et al., 2019), completing five 30-minute simulated drives over consecutive days. The original experiment took place in a medium-fidelity driving simulator comprised of a right-hand drive, mark one Audi TT car. Three ceiling-mounted, high-definition projectors provide an approximately 270 degrees forward and side view of the dynamic, unfolding driving scene onto a curved screen, with edge-blending and image warping ensuring a contiguous image is presented to participants. Side mirror displays are integrated within the original mirror housings. The rear view is displayed via a display screen placed behind the vehicle. Dashboard data was created using a bespoke application and presented on a 7-inch LCD screen, replacing the original Audi instrument cluster. A Thrustmaster T500RS force feedback steering wheel and pedal set are integrated with the existing Audi primary controls and cabin environment. Videographic data was captured of each drive, using 4 video cameras placed within the cabin of the vehicle. Each camera's position provided different vantage points of the driver sitting in the driver's seat so that observations could be made of drivers' behaviour. See figure 3-2 for an example of the 4 camera views. A total of 245 videos were captured (49 participants x 5 simulated drives).

All experimental drives included a TOR issued due to the vehicle approaching an operational design boundary limitation. These planned control transitions are predictable and, therefore, can be issued in advance, providing drivers with time to prepare themselves to resume manual driving. An additional emergency TOR was included on the fourth consecutive drive. These failure-induced control transitions are not predictable and are issued due to a critical event or ADS failure, thereby requiring the driver to resume manual driving much faster than during a planned control transition. The TORs consisted of a series of audio-visual alerts on an in-vehicle human-machine interface (HMI), which provided both text and images for visual feedback and matched spoken text and tones for audio feedback. A speech to text app was used to create a spoken, male voice alert, which was used for all spoken alerts for the HMI. The transition information displayed on the HMI varied between participant groups. For example, some groups received additional cues such as a visual countdown bar or timer and a spoken text instruction to 'check for hazards'. Other groups received more basic information (see Figure 3-1). A spoken "Manual Mode Engaged" alert was triggered as the "Manual Driving" display appears. The end of the spoken text marked the end of the automation mode and control was handed back to the driver. Drivers were given the option to

resume manual driving from the automation at any point using a verbal command.

Transition process					
Autonomous driving HMI display	Group	TOR Type	HMI Display split by group and TOR type		Manual driving HMI display
	Basic information	Controlled TOR	60 second display	10 second display	
			Prepare to drive	Resume control 	
		Emergency TOR	Resume control		
			Resume control 		
	Additional Cues	Controlled TOR	Prepare to drive	Check for hazards 	
Emergency TOR		Check for hazards 	Resume control 		

Autonomous Driving

Manual Driving

Figure 3-1 HMI display alerts split by participant group and TOR type (Burnett et al., 2019)

The focus of this study was driver behaviour during the planned control transitions from automated to manual driving. However, the failure-induced control transitions were included in coding and data analysis for relevant points to support the interpretation of results. For the purposes of video analysis, the period of ‘control transition’ was defined as: from the timestamp that marked the audio-visual “Prepare to drive” HMI alert, to the point the driver was perceived to have resumed control of the driving task. This was determined using observation of salient driver behavior, interpreted as demonstrating the driver was relaxed and in full control of all aspects of the driving task. For example, a combination of the vehicle visibly being laterally and longitudinally controlled, alongside a change in the driver’s posture, a sigh, or a change in their hand position on the steering wheel.

Coding categories were selected following initial observations of salient driver behaviour from three videos of different drivers. Categories included: changes in hand or body posture movements, such as hand(s) placed on the steering wheel, or shifting body in the seat/sitting upright; verbal or physical ‘startle’ reactions, such as a gasp or jolt of the body; and salient glance fixation locations, such as the HMI, dashboard, rear-view and side mirrors, or external to the front windscreen or left/right windows. Key points within the control transition were also coded and recorded, such as the ten second resume control countdown. This allowed timestamps of behaviour to be analysed, in accordance with standardised system events during the transition, for

example, HMI alerts and manual driving mode engagement. Ad hoc notes were also made to provide additional detail for coding categories to support and enhance coding information. For example, hand position on the steering wheel (represented as numbers on a clock face) and the type and number of NDRT being used.

Behaviour Observation Research Interactive Software (BORIS Version 7.4.7; Friard and Gamba, 2016) was used to create and store an ethogram. Videos were uploaded into BORIS and behavioural observations were coded and time stamped during the defined 'control transition period' for all driver videos. Behavioural coding from all videos was aggregated and analysed using quantitative (e.g. total count of glance or action behaviour during set time period) or qualitative observations (e.g. using multiple NDRTs at the same time). As this was an exploratory rather than evaluative study, no statistical analysis was conducted on the findings of this study. Instead the qualitative behavioural findings will be used to inform and support the development of intervention concepts, which will later be evaluated in a more structured simulator study.

3.5 Results and Discussion:

Results and discussion focus on key observations of driver behaviour during the transition of control (i.e. the sixty second time period from the issue of the TOR to the point manual driving mode was engaged). Video analysis was used to categorise glance behavior. 'Readiness to drive' was defined as the driver having both hands on the wheel and having made at least one glance towards the front windscreen. This followed the pattern of salient motor behaviour typically used to determine driver readiness within research on control transitions (e.g. Zeeb et al. 2016). The number and duration of behaviours were calculated from time stamp data extracted from BORIS. For example, where a glance at the front windscreen (taken as eyes on the road ahead), preceded a glance at the HMI, 'glance duration on the road' was calculated as: Time stamp glance (HMI) – Time stamp glance (front windscreen) = glance duration (road).

Analysis of driver behaviour suggested that use of a driver-led choice of NDRT during automation added a layer of complexity to human behaviour in control transitions, not seen in studies that use a prescribed NDRT. Janssen et al. (2019) posited that when humans are required to switch between two different tasks, the time-on-task will be affected by individual priorities. Driver behaviour provided evidence that engagement with an NDRT of personal interest to the driver increased distraction at the point of transition. Following the 'Prepare to drive' alert where the NDRT and driving tasks conflicted, drivers appeared motivated to continue interacting with the NDRT. For example, a number of drivers throughout the study continued to glance at their

NDRT after the audio alert to 'Prepare to drive'. By the fifth consecutive journey the number of drivers who continued to glance at their NDRT after this alert had been issued increased, as did the number of glances made to the NDRT (see Table 3-1).

Table 3-1 The number of drivers of glanced at their NDRT after the 'Prepare to Drive' TOR alert and those who glanced more than 5 times at their NDRT

	At least one NDRT glance	% total drivers >1 glance	≥ 5 NDRT glances	% total drivers ≥ 5
Day 1	22	45	4	8
Day 5	25	51	6	12

Additionally, there was an increase in the number of drivers who interacted with their NDRT after the TOR or where the NDRT directly or indirectly continued to provide distraction to the driver once the driving task had been resumed, impacting the time it took drivers to turn their attention fully to the driving task. For example, continuing to interact with a laptop after the TOR alert has been issued, or glancing at the phone after manual mode had resumed (see Figure 3-2 and Figure 3-3).



Figure 3-2 Driver continues to interact with laptop after 'Prepare to Drive' HMI audio TOR alert



Figure 3-3 Driver distracted by their mobile phone placed on the passenger seat after manual driving mode has resumed

Other examples included where the driver had re-organised the driving cab during automation mode, and readjusted these physical changes, either during the HMI instruction to check for hazards, or after manual mode had resumed. For example, turning off the reading light or adjusting the seat (see Figure 3-4 and Figure 3-5). Even during the emergency TOR scenario some drivers took care to 'save' where they were in their NDRT before turning their attention to the driving task. For example, placing their bookmark into the book, or closing programs on their phone, before placing the item on the passenger seat.



Figure 3-4 Driver switching internal reading light off after manual driving mode has resumed.



Figure 3-5 Driver adjusting seat during the emergency scenario TOR and audio 'resume control' alert.

Results from the original study (see Large et al., 2019; Burnett et al., 2019) showed an improvement in driver lateral and longitudinal control performance at the point of manual mode engagement over the five days. However, observed behavioural patterns suggested that this did not necessarily portray a state of driver readiness. Although drivers demonstrated a pattern of learning and applied strategy to improve performance of the control level tasks involved in the take-over, they did not demonstrate the same pattern or focus on improving tactical level task performance. This pattern became more noticeable by day 5, for example, the average time it took for drivers to place both hands on the wheel, after the 'Prepare to drive' alert, increased from 16 seconds on day 1, to 20 seconds on day 5. There were also a number of drivers who used the timer on the HMI as a countdown to cease interaction with their NDRT and switch attention to the driving task, rather than a period to carry out actions to increase performance of tactical level tasks (Figure 3-6). This behaviour suggests that drivers placed importance on controlling the steering wheel, but not on gaining situational awareness prior to the controlled transition. It also suggests that as the drivers gained familiarity with the HMI, they found ways to use it to reduce mental workload during the transition, suggesting evidence of the use of satisficing (Kaber, 2018).

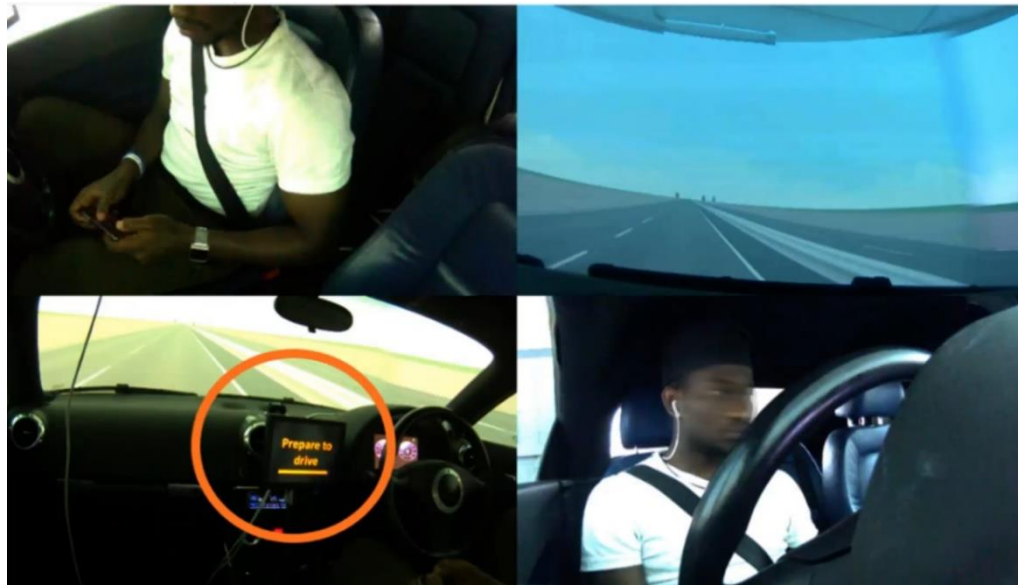


Figure 3-6 Driver using the countdown timer bar on the HMI display as a cue to disengage from using their NDRT

Driver eye glance behaviour also suggested a learning strategy. For example, the average time it took for drivers to glance at the road after a controlled TOR had been issued decreased from 7 seconds on day 1, to 4 seconds on day 5. There was also an increase in the number of drivers who had their eyes on the road at the point of transition, whereas on day 1 there was a pattern of drivers looking at the HMI at the point of manual mode engagement (Table 3-2).

Table 3-2 Number of drivers with eyes on HMI, dashboard or front windscreen at the point of manual driving mode engagement on day 1 and day 5

	HMI		Dashboard		Front windscreen		Total
	No.	%	No.	%	No.	%	
Day 1	16	33	11	22	22	45	49
Day 5	5	10	5	10	39	80	49

These factors suggest that drivers learnt the importance of looking at the road for transition performance. However, other behaviours suggested tactical tasks were not prioritised in the same way. For example, although the average time it took drivers to look at the external environment after the “Prepare to drive” alert, decreased by day 5, the time between the first and second glances at the road, increased by the same amount of time. This was because drivers turned their attention immediately back to the NDRT after the first glance at the road. There were also a number of drivers who requested take-over prior to the vehicle handing control to them, but they requested this without having looked at the road, and in some cases, without their hands on the wheel. There were also instances where drivers showed a prolonged focus on the speedometer between glances at the road. When coupled with the increased prioritisation of NDRTs, and the use of the HMI count-down to acknowledge when engagement with the NDRT needed to cease, this behaviour suggests

that 'eyes on the road' was a strategy to aid lateral driving performance, rather than the driver recognising that increasing their situation awareness was a necessary task to get them ready for manual driving. This is an important point, especially when considered with the increase in delay to placing hands on the wheel. It suggests that as driver's knowledge of the system and how to operate it increased, so did their propensity to use satisficing to aid their decision making associated with preparing to drive, even following the emergency TOR. Additionally, the use of singular, momentary glances towards the external driving environment is not enough to reflect an accurate and appropriately calibrated attention to support good situational awareness in accordance with the literature in this area. For example, higher levels of situational awareness are associated with increased attentional focus and vigilance to the external environment, so that drivers can build and maintain the information required to proactively anticipate potential future events (Fuller, 2005; Seppelt et al. 2017). Instead, a sustained pattern of appropriately time glances to important and/or relevant areas of interest in relation to the events on the road would be indicative of behaviour that would promote well calibrated situational awareness (Endsley, 2018). Noting that the definition often used to depict 'readiness to drive' has a threshold of one glance to the roadway (e.g. Zeeb et al. 2016), it is clear that this definition is not an adequate representation of driver readiness. Future definitions (that ultimately drive mechanism with which to measure driver readiness) should take into account how to measure glance behaviour that evidences appropriately calibrated attention for 'good' situational awareness.

The storyboard in Figure 3-7 illustrates these points. On day 1 (image 'a') the driver's gaze is focused on the HMI at the point of manual mode engagement, whereas on day 2 (image 'b') they made a conscious effort to focus their gaze on the road at the same point, suggesting learnt behaviour to improve transition performance. On day 5 (image 'c' and 'd'), the driver made a verbal request to resume manual driving, but their eyes were not on the road and both hands were not on the steering wheel, suggesting that the request was made ahead of the driver being 'ready'. When the driver put their hands on the steering wheel, they were positioned at the base rather than in the 10-to-2 position as they were on day 1 and 2, suggesting that the driver was feeling more relaxed about the transition to manual driving. Additionally, although the driver had their eyes on the external driving environment at the point of manual mode engagement, they immediately moved their focus to make two longer gazes at the speedometer (1 and 1.5 seconds long, with a 0.5 second glance at the front windscreen between). This again suggests that the driver was focused on controlling the operational driving tasks and placed little focus on monitoring the external environment to regain situation awareness.

- a) Day one, driver glances towards HMI at the point of manual mode engagement. Hands 10:2 position on steering wheel.



- b) Day two, driver makes long glance towards the roadway with minimal blinking at the point of manual mode engagement. Hands 10:2 position on steering wheel.



- c) Day five, driver requests early manual control while glance is towards the floor and one hand is loosely on the steering wheel.



- d) Day five, driver makes glances fixed towards speedometer immediately after manual driving mode engagement. Hands placed at the base of the steering wheel.



Figure 3-7 a) – d) presents a storyboard of driver salient behavioural adaptations from day one, day two and day five.

It can be inferred from these outputs that drivers did not recognise a need to undertake tactical level tasks during the transition phase of the take-over. The lack of engagement with the driving task, in particular monitoring of the external environment during the “Prepare to drive” phase, suggests evidence of complacency. The speed at which drivers were able to learn dynamic take-over of control level tasks, and their limited experience of uncertainty, may have led drivers to develop an unjustified satisfaction with the shared human-machine system and, therefore, failed to recognise the importance of tactical level task performance for maintaining system reliability. This observation suggests that an increased exposure to system failure scenarios through training interventions could improve transition performance. However, evidence of the use of cognitive heuristics, such as satisficing, demonstrates a desire to reduce demand on cognitive resources, which may remain even after training interventions. Employment of heuristics could reflect an increase in cognitive demand of carrying out control tasks during a transition and, therefore, a prioritisation of cognitive resources towards control rather than tactical level driving tasks. However, it could also reflect a motivation to prioritise NDRTs until the moment of manual mode engagement, which could have long term implications on drivers ‘readiness’ to drive. These findings highlight the importance of accurately defining and applying concepts such as complacency, heuristics and biases, when looking for patterns in behavioural adaptations in automated driving, particularly for the purposes of the identification and selection of effective mitigation strategies.

3.6 Conclusions and next steps

This study presented patterns of human behaviour that raise important points regarding the interaction of human drivers with vehicles with L3 ADS. Findings showed that drivers placed an emphasis on learning how to master control level driving tasks, whilst neglecting tactical level tasks, demonstrating evidence of complacency. The amount of focus drivers gave to improving tactical level tasks during the transition phase showed no evidence of change, despite performance improvements on control tasks and drivers’ exposure to a failure scenario. Instead, as drivers gained experience and knowledge of the system, they adopted cognitive strategies, such as the use of the satisficing heuristic, to allow prioritisation of NDRTs. This suggests that, at the point of manual mode engagement, drivers may not be ‘ready’ to effectively carry out tactical level tasks and resume complete control of the driving task. However, perhaps more importantly, it highlights that training or system design strategies aimed at combating the effects of complacency by increasing driver system knowledge and performance, whilst necessary, may not be a sufficient mitigation strategy for behaviours such as satisficing.

These findings have important implications for the direction of the thesis and provide foundational points to shape the design of potential behavioural interventions for experienced human drivers. Additionally,

they highlight the importance of concepts such as complacency to accurately identify and interpret patterns of behaviour, so that appropriate mitigation strategies can be found. Phase three of this thesis aims to design potential training and technological solutions to support desired driver behaviour during planned control transitions in vehicles with conditional driving automation. The findings presented here will be used to inform the direction and focus of the development of these behavioural interventions, which aim to guide desired driver behaviour and mitigate against unfavourable operational deviations. Further driving simulator studies will be conducted to evaluate the effectiveness of these interventions. These studies will also continue to investigate the impact of behavioural adaptations on driver performance following resumption of control to assess the necessity of behavioural interventions for experienced drivers.

4. Design of Driver ‘Behavioural’ Training and the ‘CHAT’ Standard Operating Procedure

4.1 Introduction

It is evident from the work carried out during the literature review and the observation study that the introduction of conditional driving automation to the driving task will place complex, new demands on the human driver. Research within the field of human factors (e.g. Kyriakidis et al., 2019) suggests that human drivers are likely to be ill-equipped to meet these new demands. Despite these anticipated performance shortcomings, empirical studies show that many drivers are highly willing to engage in immersive, non-driving related tasks (NDRTs) and activities (De Winter et al., 2014), even when the vehicle is operating at lower levels of automation and despite these new demands. This suggests that drivers are either unaware of their new role and responsibilities in future automated vehicles or actively chose to overlook them during these studies.

There is a clear disconnect between our understanding of the impact of automation on human behaviour and performance and the expectations of future vehicle technologies held by potential users. These findings suggest that the specific new skills associated with different automated driving use-cases may be unclear, difficult to imagine, and yet to be defined. Nonetheless, the evidence presented in the thesis so far points to a need to facilitate a change in drivers’ knowledge and behaviour to ensure they act appropriately at different stages and levels of vehicle automation. Indeed, it is feasible that some of the key operational skills associated with automated driving may not necessarily be particularly new or novel – and may already feature in an experienced drivers’ ‘toolbox’. However, it is their application and maintenance in the context of driving and interacting with vehicles equipped with partial and conditional driving automation that is problematic. As such, it is suggested that a prudent first step would be to change drivers’ overall behaviour and understanding of the driving task in this new context, rather than identifying and imparting a specific new skill, *per se*. Nevertheless, this clearly does not negate the need to subsequently identify and impart new and enhanced operational skills as our understanding of the role of the human driver, and their capability, within different levels of automation becomes better understood.

This chapter outlines the design and development of a proof-of-concept behavioural change training intervention focused specifically on vehicles with level 3 automated driving systems (L3 ADS), in accordance with the Society of Automotive Engineers (SAE) previously outlined (see **Error! Reference source not found.**).

4.2 Behavioural change training

4.2.1 Learning approach

Literature in the field of training design advocates that trainees should be active agents in their own learning (Bell and Kozlowski, 2008). Active learning is an effective approach in targeting higher order cognitive abilities. Driver education research has highlighted the importance of addressing areas such as driver motivation and goals, as well as developing lower-level skill-based behaviours (Hatakka et al., 2002; Fylan, 2017). The ability of a learner to recognise how well they are performing and to judge when they are likely to be accurate or when they are likely to make a mistake (known as metacognition (Bell and Kozlowski, 2008)), alongside the constructive framing of errors and their consequences, are critical to achieving active learning (Krampell et al., 2020). Therefore, behavioural change theories (Fylan, 2017) were drawn upon to target driver motivation to perform desired behaviours, using the following four principles:

- **Shaping Knowledge:** Providing the driver with a better understanding of what causes them to behave in a certain way, and the knowledge of how to perform the target behaviour.
- **Natural Consequences:** Highlighting the consequences of undesirable behaviours, including the social, environmental and emotional consequences, and making these consequences seem more real, for example, by making learners realise that they would regret failing to change their target behaviour.
- **Comparison of Outcomes:** Considering the outcomes of performing desired behaviours, or not, and understanding why the target behaviour is a good thing to do.
- **Antecedents:** Understanding the triggers for the undesirable behaviours, taking practical steps to avoid those triggers, and changing the physical and social environment to make it less likely that drivers will perform these behaviours and more likely they will carry out the desired behaviour.

These four principles were applied in the context of level 3 ADS to devise a behavioural training intervention that could be delivered to experienced drivers. The training was designed to be self-administered, with visual lesson content delivered via a PowerPoint presentation and including voiceover instructional/expert commentary. A professional actor narrated the presentation to increase authenticity, and reduce any bias associated with the same experimenter both voicing the presentation and conducting the study itself. The overall aims of the behavioural training were to:

1. improve drivers' understanding of vehicle automation;
2. outline drivers' roles and responsibilities within level 3 automation; and
3. provide best practice guidance for driving and interacting with a level 3 automated vehicle, including communication of a specific operating procedure.

The behavioural training intervention therefore aligns with Degani and Weiner's (1994) 3Ps of flight deck operations in the aviation domain (i.e. Philosophy, Policy and Procedures), and Casner and Hutchins' (2019) adaptation of them for use in vehicle automation. The aim was to meet the following learning objectives through the intervention:

- Drivers will have a better understanding of Vehicle Automation, including the SAE Levels of Automation and '**Automation Philosophy**': the reasons for/benefits of automation, in particular how vehicle automation 'augments' the driver's role rather than removes it.
- Drivers will have a better understanding of the capabilities and limitations of a level 3 vehicle, understand the concept of 'shared control' and a clear mental model of their role and responsibility as a driver. This will include a focus on an '**Automation Policy**': an understanding of when it is and is not appropriate to use automation.
- Drivers will have a better understanding of some of the consequences of vehicle automation on driver behaviour. Drivers will have knowledge and understanding of '**Automation Procedures**': the desired way to engage with a level 3 vehicle. This will focus on the use of non-driving related tasks during automation and awareness of key tactical level controls e.g. monitoring, attention, situation awareness etc., required when interacting with a level 3 automated vehicle.

The behavioural training also focussed on the control, tactical and strategic subtasks of the driving task, with an emphasis on the spatiotemporal elements associated with situation monitoring for each of these levels. The aim of this was to represent the goals of situation monitoring at each level of the driving task in terms of task importance and urgency in relation to the different system states; their role in prediction and decision-making at each level and the potential consequences of neglecting situation monitoring during automated driving (see Figure 4-1).

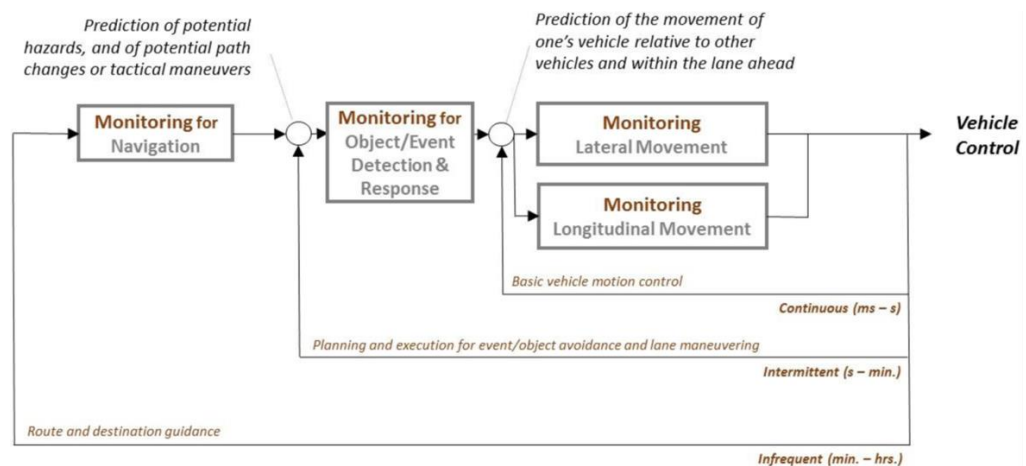


Figure 4-1 Monitoring inherent to multi-level control in driving (Merat, 2019)

- **Control monitoring goals ('Continuous'):** These goals involve controlling basic vehicle operations and managing the vehicle in relation to other vehicles on the road within the lane ahead. This monitoring of lateral and longitudinal control is delegated to the automated systems during automated driving but is required to be resumed by the human driver during and following a transition to manual driving mode.
- **Tactical monitoring goals ('Intermittent'):** These goals involve managing the vehicle in relation to planning and executing events along the journey, including object avoidance. This monitoring should be carried out intermittently by the human driver during automated driving, as good practice, to support the prediction of system limitations and controlled (or emergency) takeovers. Following a takeover request (TOR) and during the transition phase, this monitoring should be conducted with greater frequency and urgency to support the initial takeover and facilitate proactive driver decision making in relation to the planning and execution of the next event/manoeuvre. For example, to plan for a junction exit from a motorway the driver needs to be aware of the vehicles surrounding them on the road, including in their blind spot, **before** taking physical control of the vehicle, so that they can effectively navigate any path changes or manoeuvres. For instance, to avoid surrounding objects and potential hazards in time to safely take their exit.
- **Strategic goals ('Infrequent'):** These goals refer to managing the driving task in relation to the desired destination. It requires infrequent monitoring and attendance to salient and relevant cues in the driving environment associated with navigation. This monitoring should be maintained by the driver through the whole journey, regardless of driving mode, to ensure, for example, that spatial awareness is maintained and that the chartered route is the best, fastest and most appropriate to meet the driver's needs.

4.2.2 Mental model formation

To support the development of an accurate and complete mental model, trainees were given an overview of the SAE Levels of Automation, detailing their roles and responsibilities as the driver at each level (Figure 4-2). In addition, they were provided with a precise explanation of the terms 'Automated' and 'Autonomous' vehicles (Figure 4-3), in the context of commercially-available vehicles and likely future development. Drivers were subsequently given more detailed information relating to SAE level 3 vehicles and their limitations and capabilities (Figure 4-4), and were reminded of the aspects that were controlled by the vehicle during automated driving and those of which the driver needed to retain an awareness.

Levels of Automation

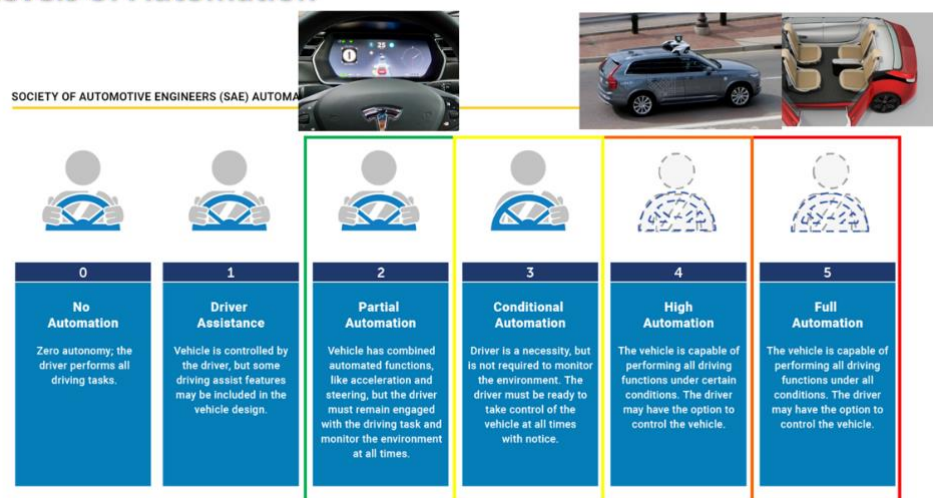


Figure 4-2 Training screenshot showing the SAE Levels of Vehicle Automation and the roles and responsibilities of the human driver at each level (Adapted from SAE International, 2016)

Automated vs Autonomous vehicles

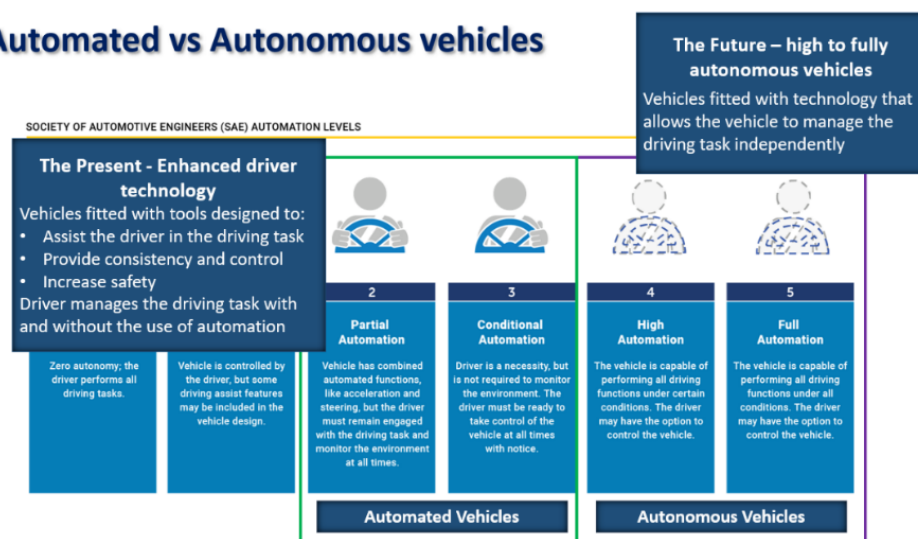


Figure 4-3 Training screenshot explaining the definitions of 'Automated' and 'Autonomous' vehicles, showing blue dialogue boxes used during the presentation (Adapted from SAE International, 2016)

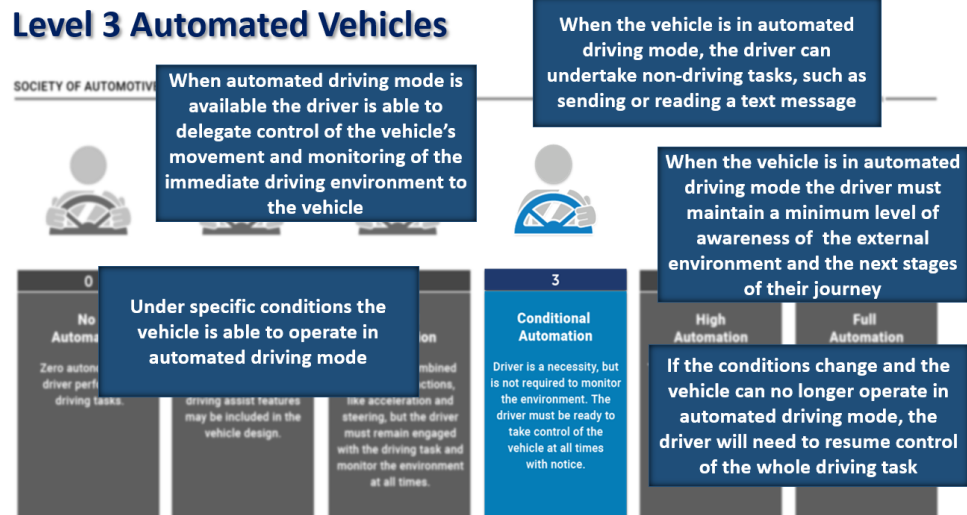


Figure 4-4 Training screenshot explaining SAE Level 3 Automation, showing blue dialogue boxes used during the presentation (Adapted from SAE International, 2016)

Specifically, drivers were told that they were responsible for the actions of the vehicle during all parts of the journey, but that when they engaged automated driving mode, the vehicle was capable of managing the control level monitoring goals of the driving task, thereby increasing the comfort of the driver. However, drivers were reminded that a level 3 automated vehicle is only able to operate in automated driving mode under certain conditions. These conditions would vary based on the capability of the specific make and model of the vehicle, but could include the following limitations:

- type of road e.g. use of dual carriageway/motorway;
- weather condition; and
- amount of daylight.

Drivers were told that if the vehicle approached a limitation where automated driving mode could no longer perform, it would issue a TOR to the driver, requiring them to take manual control of the vehicle (Figure 4-5). Moreover, they were instructed that although automated driving mode allows the driver to undertake activities that are not related to driving, such as using a mobile phone, it is imperative that they, the driver, maintain an awareness of the tactical and strategic goals for the vehicle so as to:

1. predict when the vehicle might issue a TOR;
2. maintain an awareness of where the vehicle is on the journey and what next steps are required should a TOR be issued; and

3. allow the driver to focus on the short-term goals when managing the takeover.

Automated driving mode

- ⚠ Automated driving mode can only operate under certain conditions
- ⚠ When those limits are approached the car will issue a request for the driver to take control of the full driving task.
- ⚠ Drivers must maintain a minimal level of awareness during automated driving mode so that they can smoothly take back full control of the driving task.



Figure 4-5 Training screenshot discussing automated driving in relation to Operational Design Domain (ODD) limitations, the role and responsibility of drivers and situation monitoring of tactical and strategic tasks to promote on-the-loop engagement

As the literature has shown, mental models allow people to derive appropriate behaviour for interactions with objects, events and situations, through the provision of explanatory knowledge on the way things work (Norman, 2013). Mental models make learning easier, but if designers or training providers do not offer appropriate mental models, people are likely to create their own, inappropriate ones (Norman, 2013). In the current context, these could lead to the misuse or incorrect use of driving automation (Beggiano and Krems, 2013; Casner and Hutchins, 2019; Merat et al., 2019). Therefore, the first half of the training was designed to encourage the development of accurate, and complete, mental models of vehicle automation and the functionality, limitations and capability of level 3 automated vehicles. This was achieved through the use of appropriate wording, explaining why particular actions are required, as well as what the actions are (Wickens et al., 2014).

Although explanatory knowledge provides meaning and structure that is fundamental to human performance, retrieval of this knowledge from long-term memory stores involve mental resources that are not ideal for tasks necessitating rapid or smooth action (Norman, 1988). Therefore, whilst the development of a good mental model is necessary for the driver to learn appropriate interaction with the vehicle, it is unlikely to be sufficient for supporting the actions required by drivers to calibrate their situation awareness (SA) and attention in a timely manner (Carsten and Martens, 2019). Rote learning of operating procedures and checklists

have been successfully used in aviation as a necessary and efficient way to facilitate the fast recall of information by pilots in emergency scenarios (Norman, 1988; Degani and Wiener, 1997). Yet, reliance on arbitrary knowledge alone makes learning difficult and does not enable the human to apply action in meaningful ways that allow for subtle adjustments to nuanced scenarios, as it does not teach the reasons for the action (Norman, 1988). As such, our behavioural training was designed with the aim of combining the utility and efficiency of a clear, concise operating procedure that could be used to rapidly recall the sequence of actions in a timely manner. This was presented to participants through a meaningful and sensible structure, to simplify and organise the information, supporting them in digesting and understanding the facts, which can aid the memory task. It is anticipated that the procedure could be automatised over time, akin to the ‘mirror, signal, manoeuvre’ (position, speed, look) routine traditionally instilled as part of manual driver training (Driver and Vehicle Standards Agency (DVSA), 2024).

The use of HMI design as a way to support desired driver behaviour (as a countermeasure to some of the issues discussed above), and reinforce operating procedures (Carsten and Martens, 2019) form important longer-term goals for future directions of this research (following on from the proof-of-concept phase). Therefore, our design approach to the operating procedure also took ecological interface design principles (Rasmussen and Vicente, 1989) into consideration and aimed to promote a reflexive response demanded by the dynamic context of driving using rule-based behaviour to structure the application of skill-based behaviour – principally: if- my vehicle issues a TOR, then- I need to follow the ‘CHAT’ operating procedure (see below).

4.2.3 The ‘CHAT’ Standard Operating Procedure

The second part of the behavioural training introduced the ‘CHAT’ procedure (see Figure 4-6). CHAT is a standardised operating procedure (SOP) designed to motivate and support drivers in remembering and applying explanatory knowledge acquired about level 3 automated vehicles. The aim was to provide drivers with a simple and efficient way of remembering a specific sequence of actions that they must take following a TOR from the system, thereby encouraging the transition to an ‘on the loop state’ before physically resuming manual control of the vehicle.



Figure 4-6 Training screenshot showing overview of the 'CHAT' standard operating procedure

Semantically, the acronym CHAT aims to represent the verbal or non-verbal communication necessary for effective collaboration on a shared task. This can be likened to the scenario of a manager directing employees through a critical incident; to take effective and appropriate action, he/she would first need to communicate with employees on front line operations to acquire all of the relevant information. The letters represent the actions: CHCheck (CH-), Assess (-A-) and Takeover (-T). This aims to guide the driver in making the appropriate checks and assessment of their internal and external environments prior to taking over operational control of the vehicle (see Appendix B for a transcript of an extract from the training script used to describe CHAT during the training). For example, during automated driving, drivers may have adjusted the driver cab, which need to be reset. Additionally, drivers need to check the external environment surrounding the vehicle, including their blind spot and via their mirrors. They will need to assess their position and speed, as well as predicting the movement of other road users in relation to any manoeuvre needed following the takeover. In effect, this procedure is designed to allow the driver to recombobulate, by promoting an 'on the loop' driver state during the transition period. The procedure aims to encourage drivers to increase their situation monitoring and improve their SA prior to physically taking control of the vehicle. The expectation is that this will ameliorate the transition for drivers from OOTL to in the loop, which will, in turn, support reasoned decision making for tactical or strategic tasks following takeover.

'CHAT' is designed in view of several guiding principles (see Wickens et al., 2014) to aid learning, memory and understanding in relation to the application of operating procedure it represents:

- **Standardisation** – the use of CHAT provides a way to standardise this operating procedure across the field of automation (akin to ‘mirror, signal, manoeuvre’ (position, speed, look) (MSM/PSL) in current driver training).
- **Use of a memory aid** – the use of an acronym constrains the representative action words to the digraph and phonetic sounds: ‘CH’, ‘A’, and ‘T’ to provide recognition opportunity from ‘bottom up’ information (e.g. from an HMI) to aid ‘top down’ recall of information from memory.
- **Use of memorable characteristics**: the semantic association with a request for communication (from the system partner), use of a concrete (rather than abstract) word and no technical jargon; and representation of an organised set of information e.g. the order of the letter sounds represents the required sequence of the operating procedure, the timing of the ‘CHAT’ instruction can also be scheduled via a HMI to promote timely scheduling of the procedure with the driver.

Following an initial explanation of the CHAT procedure, a learning strategy called proactive observation (Castro et al., 2016) was used, combining the approaches of meta-cognition and expert (road) commentary. The aim of this was to motivate trainees to adopt the CHAT behaviours by drawing their attention to potential errors that can be made through inattention. Trainees were presented with a bird’s eye view of a takeover scenario (Figure 4-7) and were instructed to actively scan the road scene, applying the principles of the CHAT procedure to establish the ‘Checks’ and ‘Assessments’ they would need to make if they were the driver in the ‘ego’ blue vehicle. Expert commentary then guided them through the task, providing immediate feedback and highlighting the importance and utility of the CHAT procedure. The commentary also indirectly highlighted the consequences of *not* carrying out this procedure ahead of the takeover given spatio-temporal constraints in this example. Specifically, the commentary accompanying the images in Figure 4-7 was scripted to:

1. Draw trainees’ attention to the ‘ego’ blue car (centre of shaded red circle) and ask them to imagine what they would need to do if they were the driver and had just been issued a takeover request.
2. Identify the ‘Checks’ required in relation to the CHAT procedure, drawing attention to the vehicles in front, behind and in the blue car’s blind spot and providing immediate feedback to the trainee.
3. Explain the ‘Assess’ element of the CHAT procedure, drawing trainees’ attention to each of the vehicles surrounding the blue car in turn, building upon the interrelationships between each of them and the movement of the blue ‘ego’ car. Through this

process, emphasis was placed on the necessity of carrying out the CHAT process **before** 'Takeover' of the physical control of the vehicle.

Resuming the whole driving task

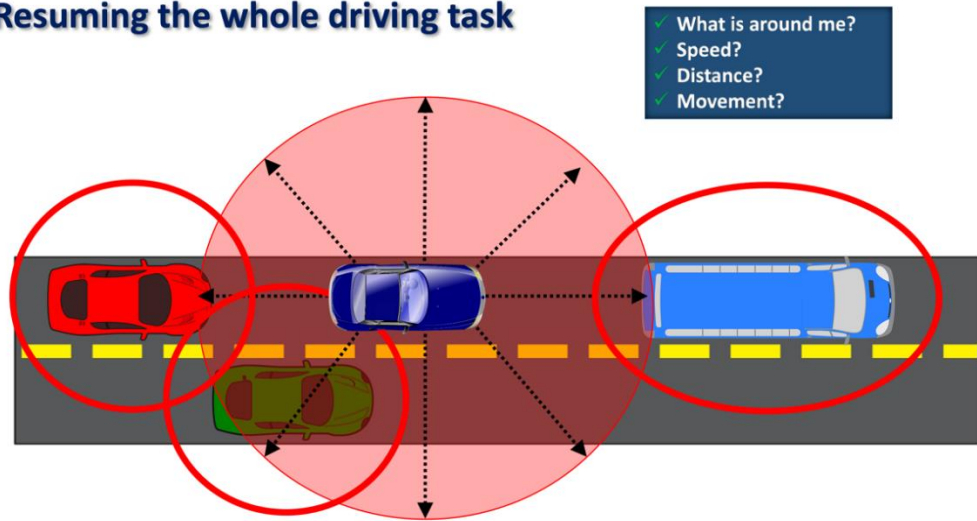


Figure 4-7 Training screenshot of interactive slide used with expert commentary for proactive observation of the CHAT procedure. Blue 'ego' vehicle appears at the centre of shaded red circle

4.3 Operational training

To assess the effectiveness of the CHAT behavioural change training compared to current "training", a 'user manual' was also created (extract in Figure 4-8). This was based on the existing manual provided by a commercially available vehicle fitted with level 2 automated driving systems (ADS) (Tesla Company, 2019). The user manual mirrored the style and number of warnings and advisory notices related to engaging with the automated system. However, it was updated to reflect both driver requirements at level 3 automation and the specific operational features of the simulated vehicle (see Appendix C for full document). The user manual provided an overview of how the level 3 vehicle worked, detailed the automated features that were fitted in the vehicle ('Autosteer', 'Advanced Cruise Control' and 'Advanced roadway monitoring'), and explained the capabilities and limitations of each feature. There was also a separate section on the limitations of 'Automated driving mode', which detailed numerous warnings relating to the fall-back requirements for the human when the vehicle is in automated driving mode. Finally, there was a section on how to operate automated driving mode, which mirrored what the driver would need to do during the experimental drive. The intention in creating the user manual was to provide a baseline (reflecting current practice) against which to evaluate the performance and effectiveness of the behavioural change training. The user manual was subsequently referred to as 'Operational training'.



When automated driving is engaged and the vehicle approaches a condition where automated driving will no longer be available for use, the vehicle will issue advanced notice for you to resume full, manual control of the vehicle.

Limitations

Many factors can impact the availability of the Automated driving mode within Autocar3, causing this mode to be disabled in certain conditions.

These include (but are not limited to):

- Poor visibility (due to heavy rain, snow, fog, sunlight etc.).
- Damage or obstructions caused by mud, ice, snow, etc.
- Interference or obstruction by object(s), mounted onto the vehicle (such as a bike rack).
- Obstruction caused by applying excessive paint or adhesive products (such as wraps, stickers, rubber coating, etc.) onto the vehicle.
- Narrow or winding roads.
- A damaged or misaligned bumper.
- Interference from other equipment that generates ultrasonic waves.
- Extremely hot or cold temperatures.



Warning: The list above does not represent an exhaustive list of situations where automated driving is not available for use. Never depend on these components to keep you safe. It is the driver's responsibility to stay alert, drive safely, and be ready to take full control of the vehicle at all times.

Using Automated driving mode

Engaging Automated Driving mode



You must ensure you are in the correct lane for automated driving before activating automated driving mode.

To use Automated driving, you must be driving at least 30 mph.

All alerts and information in relation to the driving modes and takeover requests will be displayed on the in-vehicle display screen, which is located in the area on the diagram below:



Figure 4-8 Page taken from the user manual created for 'Operational training' (full details in Appendix C)

4.4 Conclusions and Future Directions

As described in this chapter, the CHAT behavioural training module focuses on supporting and guiding desired driver behaviour in vehicles with L3 ADS through:

- The formation of adequate and correct driver mental models of automation, including the *automation policies, principles and procedures* associated with the related level of automation;
- Raising driver awareness and understanding of the *impact of vehicle automation on driver behaviour* and the *consequences* associated with being out-of-the-loop (OOTL) of control at each level of the driving task; and

- Establishing a *standardised operating procedure* (CHAT: Check, Assess, Takeover) to specify and communicate a clear sequence of tasks required by drivers to safely and effectively transition in and out of the loop of control during automated and non-automated driving.

This behavioural training intervention will be evaluated using a driving-based simulator study. The aim of this study, outlined in chapter five, is to assess the effectiveness of this proof-of-concept behavioural training. To do this, the training module will be compared to the user manual also described in this chapter. User manual training represents the current, typical training solution used for new, commercially available vehicle technologies. Therefore, this form of passive training was selected as a suitable comparison by which to evaluate the effectiveness of the CHAT training. Effectiveness was defined in terms of the driving control loop – that is to say, both the physical control and situation monitoring were considered in terms of driver performance. By including ‘driver monitoring performance’ in the analysis, the aim was to identify instances of driver ‘on’ versus ‘out’ of the loop states to use as a tool to investigate driver behaviour during automated driving, as well as at, and shortly after, the point of transition (Merat et al., 2019). A secondary aim of seeking to establish ‘proof-of-concept’ for the behavioural training was to identify areas for consideration in future designs of in-vehicle technological solutions. Therefore, best practice interface design principles were also considered in the training approach and analysis of the study results.

5. Driver Training Driving Simulator Study

5.1 Introduction

This chapter describes a driving simulator-based study carried out as part of the thesis to evaluate a proof of concept behavioural training module developed in chapter four. A detailed description of the experimental design and methodological approach are provided, along with the findings and overall conclusions of the study and how it links to further developments within the body of work.

5.2 Overview and Aims

The aim of this driving simulator study was to evaluate and validate the behavioural training intervention through a between-subjects study. It was compared to a more traditional (Operational) approach – effectively a printed user manual, as might be expected with current vehicle technologies. Following the training, participants were asked to undertake a drive in the simulator, which involved an episode of automated driving and a planned takeover request (TOR). Subjective measures of trust, situation awareness and workload were captured after the drive. A post-drive, self-reflective interview was also carried out with participants, while reviewing a split-screen video of the drive. Subsequent quantitative analyses explored drivers' visual behaviour, and their driving performance following the transition of control.

The goals of the behavioural training were to:

1. improve drivers' understanding of vehicle automation;
2. outline drivers' roles and responsibilities with level 3 automation;
and
3. provide best practice guidance to driving, and interacting with, a level 3 automated vehicle.

Key metrics related to driver behaviour and performance, such as checks conducted to the external environment; the time taken to fully disengage from using non-driving related devices; and vehicular control following the transition to manual driving were used to compare and evaluate the effectiveness of the behavioural training module in comparison to the operational training approach.

5.3 Methodology and approach

5.3.1 Experimental drive

The study took place in a medium-fidelity, fixed-base driving simulator at the University of Nottingham (Figure 5-1). The simulator comprises a right-hand drive Audi TT car positioned within a curved screen, affording 270 degrees forward and contiguous side view of the driving scene via three overhead high-definition projectors, together with rear

and side mirror displays. A Thrustmaster T500RS force feedback steering wheel and pedal set are integrated faithfully with the existing Audi primary controls, with the dashboard created using a bespoke application and presented on a 7-inch LCD screen, replacing the original Audi instrument cluster. Four video cameras were strategically located within the vehicle cab to record participants' behaviour from different aspects. Each camera related to a different split-screen view as shown in Figure 5-4.



Figure 5-1 University of Nottingham Human Factors Research Group Driving Simulator

The simulated driving environment was created using Systems Technology, Inc. STISIM Drive software (version 3) (<https://stisimdrive.com/>) and was similar to that used in the longitudinal driving study, the data of which were used for the observation study carried out in chapter three (Large et al., 2019; Burnett et al., 2019). However, there were a number of additional exit roads and related signage along the dual carriageway to allow for a junction exit scenario to be included in the experimental drive. The environment was designed to replicate a short commute-style journey lasting approximately twenty minutes in total. Participants began in a residential location, which was described to them as their home, then they progressed through a rural setting, and joined a UK two-lane dual carriageway. The dual carriageway supported the use of an SAE level 3 automated driving system (ADS). As such, drivers were able to relinquish the physical primary control actions, i.e. steering, acceleration and braking. Midway along the dual carriageway, participants received a TOR asking them to resume manual control as part of a planned handover.

As part of the briefing provided ahead of the experimental drive, participants were instructed that they were required to be in the right-hand lane (lane 2) in order to engage the automation and were given notice that a TOR would occur at some point during the automated drive, and they would then be required to resume control of the vehicle and prepare to leave the dual carriageway at the next exit. Participants

were provided with a brief overview of the timings and HMI display text for the transition and reminded that they had the option to override the 50 second 'Prepare to drive' portion of the transition via a button press, which would move the HMI to the 10 second 'Resume control' countdown. Participants were instructed that, following re-engagement of manual driving, they were to move into the left-hand lane (lane 1) and take the next available exit. This involved negotiating with other road users to find a suitable location to move into lane 1 before exiting the road. All roads involved were populated with moderate to high levels of traffic, authentic road signage, and geo-typical roadside artefacts and terrain. These features served to both increase ecological validity and provide cues within the external driving environment that, if attended to, could be drawn upon by participants during the post-drive interview.

Prior to attending, participants were asked to consider activities that they might like or expect to do on such a journey in an 'automated' vehicle and bring with them any objects or devices (i.e. their own belongings) that would enable these. No restrictions were placed on what participants could do to avoid pre-empting or influencing their expectations and behaviour.

To support drivers, a rudimentary HMI was installed in the centre console of the vehicle. This provided basic system feedback throughout the drive, delivered via an interactive PowerPoint presentation controlled remotely by the researcher. System feedback included information, such as whether automation was available, the status of the automated control, and guidance during the handover. Feedback was provided both visually (in the form of text-based messages and images) and audibly (by tones and spoken messages). Participants were alerted of any updates to the information presented using a tone. All participants (i.e. both groups) received the same information and notifications during the study.

5.3.2 Planned control transition

Participants began by driving manually, i.e. they were responsible for all primary control actions. SAE level 3 automated control was made available to drivers only when they joined the dual carriageway, which occurred approximately 5 minutes into the journey. Participants were notified of automation availability by means of an audio-visual text-based message presented on the in-vehicle HMI, with accompanying spoken audio. They engaged the automation via a button press and were only able to engage the automation once they had moved into the right-hand (automation) lane (lane 2) of the dual carriageway. Participants were told that during periods of automation, they were permitted to engage in NDRTs as they saw fit – no restrictions were applied other than making drivers aware at the start of the study that they would be required to resume manual control and leave the dual

carriageway at the next exit and would be given 60 seconds' warning to prepare themselves. In practice, the planned takeover request was issued approximately 10 minutes into the automated driving mode. This occurred via an audio-visual 'Prepare to drive' alert on the HMI display. This alert was displayed for 50 seconds, with the remaining time indicated as a countdown timer on the HMI display. At the end of 50 seconds, the driver received a further audio-visual alert on the HMI to 'Resume control'; this included a 10 second count down, following which manual driving mode resumed. The driver was able to manually activate the 10 second 'Resume control' countdown at any point during the 'Prepare to drive' notification period. Figure 5-2 provides an overview of the experimental drive and HMI displays.

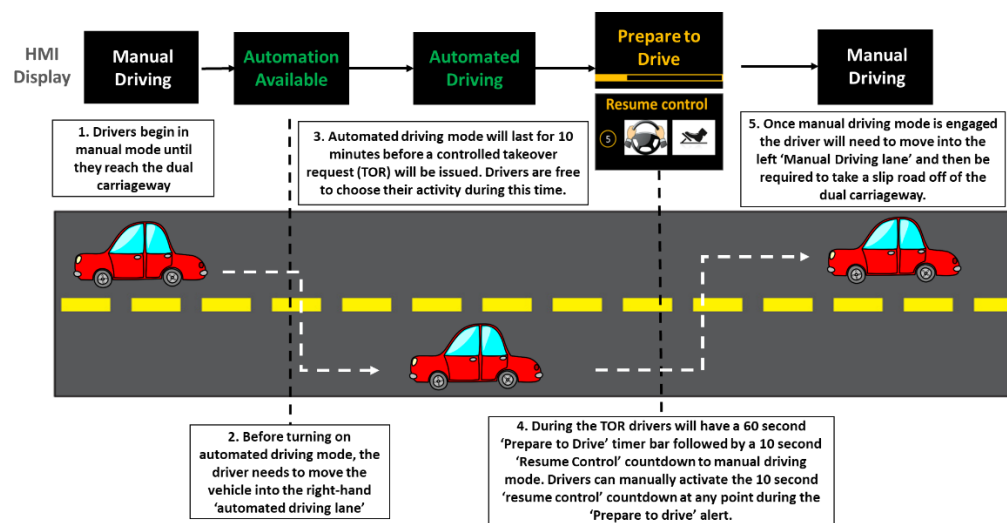


Figure 5-2 Schematic showing experimental drive and associated HMI displays

5.3.3 Participants

Twenty-five participants were recruited to take part in the study (20 male, 5 female; mean age: 35 years; range: 21-59 years). All participants were experienced drivers and were required to hold a UK driver's licence and have more than three years of driving experience and drive regularly (at least 2 to 3 days a week), including experience of dual carriageway or motorway driving (in practice, participants' mean number of years driving was 14, and their annual mileage ranged from 5,680-26,000). Participants primarily comprised employees and postgraduate students at the University of Nottingham and were recruited by means of advertisements placed around the University of Nottingham campus and sent via email. Unfortunately, one of the participants displayed some atypical behaviours (they later admitted to "playing with" the driving simulator), and their data and responses were subsequently removed prior to analysis. This resulted in approximately half the participants ($n=11$) receiving Behavioural training and the remaining participants ($n=13$) provided with Operational training prior to driving in the simulator. Participants were in attendance for

approximately one hour and were reimbursed with £10 in shopping vouchers as compensation for taking part. The Faculty of Engineering's Ethics Review Committee approved the study procedure.

5.3.4 Training

Prior to completing the simulated drive, all participants completed a 15-minute, self-directed training session. Participants either completed the CHAT behavioural training module or the Operational training. Full details are described in chapter four, a summary comparison of the training approaches can be found in Table 5-1.

5.3.4.1 CHAT Behavioural Training

Participants in the Behavioural group completed the Behavioural 'CHAT' training as outlined in Chapter four. This training was in the form of a PowerPoint presentation with audio commentary voiced by a professional actor.

5.3.4.2 Operational Training

Participants in the Operational group were provided with a document adapted from a user-manual for a commercially available vehicle fitted with level 2 automated driving system (ADS) to represent a level 3 ADS (Appendix C). The user-manual training mirrored the style and number of warnings and advisory notices related to engaging with the automated system but was updated to reflect both driver requirements for a vehicle equipped with level 3 ADS and the operational features of the simulated vehicle.

Table 5-1 Comparison of training approaches

	Behavioural Training	Operational Training
Format	PowerPoint presentation	Printed, written document
Authenticity	Commonly used training approach	Equivalent to current user manual
Content	Behavioural aspects of interacting with the automated system	Functional aspects of interacting with the automated system
Instruction to participant	<i>"Please work through the slides. Please wait for the audio to finish on each slide and then move onto the next one. The presentation should last for 15 minutes, and I will return after this."</i>	<i>"Please read through the information within this document at your own pace. Take your time to read through the information carefully, and I will return in 15 minutes."</i>
Trainer involvement	None	None
Approach	Self-administered – participant required to move through the presentation at own pace	Self-administered – participant required to read through document at own pace
Duration	15 minutes	15 minutes
Follow-up practice drive	Yes	Yes

5.3.5 Test Drive

Prior to completing the experimental drive, all participants took part in a 'test drive' scenario, lasting around 5 minutes. This allowed drivers to practise the vehicle controls, including a transition to and from automated driving and gain knowledge of the HMI display and other key features for the drive. Evidence from the previous study (Large et al., 2019; Burnett et al., 2019) showed that drivers significantly improved their lateral and longitudinal performance at the point of transition to manual mode between the first and second drives. Therefore, this test drive served to allow drivers the chance to practice the operational skills involved in the transition and offer an opportunity to validate previous behavioural adaptations (Shaw et al., 2020a) during the transition period. No other vehicles were present in the test drive scenario. Drivers began in manual mode on a dual carriageway and practised the following manoeuvres:

- changing lanes;
- engaging automation;
- receiving and responding to a takeover request; and
- transitioning from 'Automated Driving' mode to 'Manual Driving' mode.

5.3.6 Semi-Structured Interview

Following the drive, all participants underwent a semi-structured interview using the recording of their simulated drive to guide when to ask questions. A selection of pre-determined questions based on those used in the 'Deepening' stage of the Critical Decision Method (CDM) interview (Crandall, Klein and Hoffman, 2006) were used to probe the behaviour and decision making of the driver at key points along the journey, as outlined in the timeline in Figure 5-2.

CDM interviews are used to elicit information about cognitive functions such as decision making, planning and sense making within a specific event. As such, the CDM interview method offered utility as a tool to probe the cognitive processes behind the participants' behaviour and decision making while reviewing the event they experienced during the experimental drive. The aim of the post study interview and critical decision method probes was to gain a deeper and contextualised understanding of the cues, information, goals, priorities, mental models, judgements and decision making of the drivers, specifically to inform the future development of the CHAT approach and HMI design.

CDM interviews rely on retrospection and are therefore vulnerable to potential errors relating to memory. Therefore, the video recording of

the simulated drive was used to support the interview. The video, consisting of four camera views, and key events identified through the timeline construction (Figure 5-3) were reviewed by the interviewer and driver. Using the timeline, the video was paused at key points to ask various questions based on the CDM ‘deepening’ probe questions. The video recording also offered the opportunity to rewind and review key events where required.

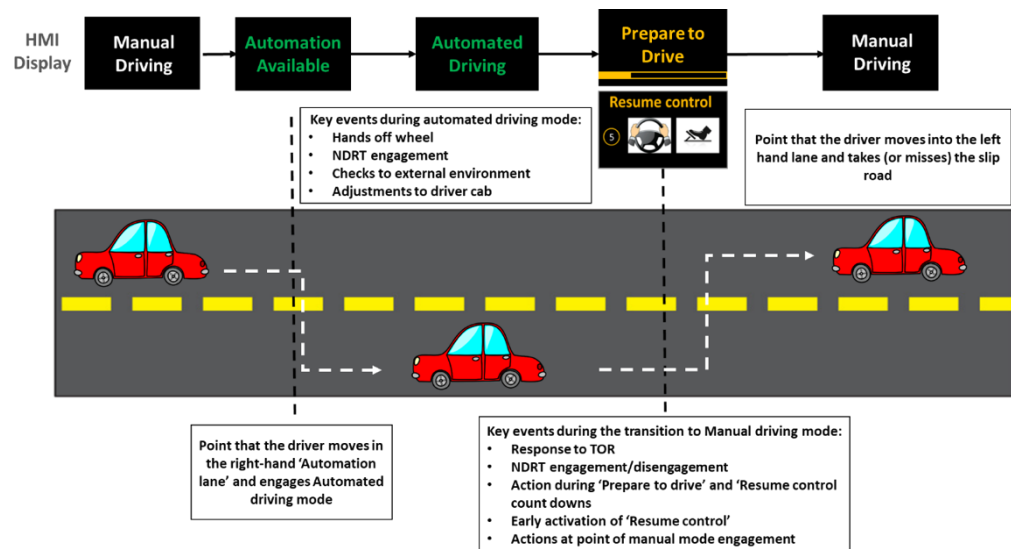


Figure 5-3 Timeline of key action points to review with the driver during the post-drive interview

Using the timeline and the video of the experimental drive as a guide, questions were selected from the CDM interview probe questions (Appendix D) to encourage the participant to elaborate on their account of what happened at each segment of the drive. The aim was to elicit the cues and information available to and used by the driver during the drive as well as what they did not have available to them that may have helped. Question choices were adapted to suit the events of the drive and in response to the behaviour observed in the video by the driver and/or the interviewer.

5.3.7 Measures

Prior to taking part in the driver training, participants completed the Total Trust in Automation Questionnaire (TTAQ) (Gold et al., 2015b). Following the training, participants were asked to complete a 5-point subjective ratings scale questionnaire adapted from a classroom engagement study (Wang et al., 2014), to evaluate the training (Training Evaluation Questionnaire, TEQ) (see Appendix A). After completion of the experimental drive, the participant repeated completion of the Total Trust in Automation questionnaire (TTAQ) (Gold et al., 2015b); the Situational Awareness Rating technique (SART) scale (Taylor, 2017) and the NASA-TLX workload questionnaire (Hart and Staveland, 1988).

Driving performance measures were captured by the STISIM simulation software. These were used to calculate key driving performance indicators associated with the takeover of control, including longitudinal and lateral acceleration, standard deviation of lane position and steering reversal rate, aiming to identify the smoothness of the resumption of control and lane change manoeuvre. In addition, the minimum time to collision (TTC) was determined to provide an indication of how close participants were to other vehicles in their immediate surroundings.

Behaviour Observation Research Interactive Software (BORIS Version 7.4.7; Friard and Gamba, 2016) was used to create an ethogram used to code driver behaviours during the experimental drive. Frame-by-frame coding of behavioural observations was conducted from the split-screen video recordings of the experimental drive; cameras were positioned so that the internal and external environment could be observed (see Figure 5-4 for examples). Observations were analysed to explore and compare behaviour between the Behavioural and Operational groups during: automated driving mode; the transition period, defined as between the TOR and the point that manual driving mode re-engaged; and during the lane change manoeuvre after drivers had resumed manual control. Behavioural measures included (list not exhaustive) eye glance direction and frequency (to mirrors, external and internal sources), engagement with NDRT, hands on/off wheel, and feet or body adjustments.



Figure 5-4 Example of split-screen video recordings used for analysis, showing participant engaged in NDRT during automation (top) and following the request to resume manual control (bottom)

5.3.8 Statistical Testing

The key goal of the training was to improve the behaviour of drivers during automated driving, and after they resumed manual control, particularly with respect to their ongoing responsibility towards driving. Therefore, a range of recognised subjective rating scales were employed to evaluate what drivers thought about their own behaviour, (for example, trust, situation awareness and workload), and how this translated into tangible benefits. Descriptive statistics were used to compare participant behaviour and identify trends. Independent t-tests were employed to determine if there was a significant difference between the means of two groups and draw inferences on the likelihood of application to a wider population. Fisher's Exact tests were used to analyse categorical data, for example, comparing lane change mirror glance behaviour between the two training conditions.

5.4 Results and Analysis

Differences in results between drivers receiving Behavioural training and those receiving Operational training were tested for statistical significance using *t*-tests and Fisher's exact test. Statistical significance is reported at $p < .05$.

5.4.1 Subjective Measures

5.4.1.1 Training Evaluation

Participants were asked to rate the training they had received (Behavioural or Operational) using the training evaluation questionnaire (TEQ) adapted from Wang et al. (2014). Ratings were captured using 5-point Likert scales, where 1 indicates strongly disagree, and 5 indicates strongly agree. Individual scales were subsequently grouped into engagement, usefulness and attitude for analysis, in line with previous applications (Lawson et al., 2019). Ratings were generally high (above the scale mid-point) for all attributes and amongst both cohorts (Figure 5-5). Notably, there were no significant differences between ratings made by drivers in the Behavioural and the Operational groups, suggesting that both approaches were comparable from the perspective of the participants' experiences.

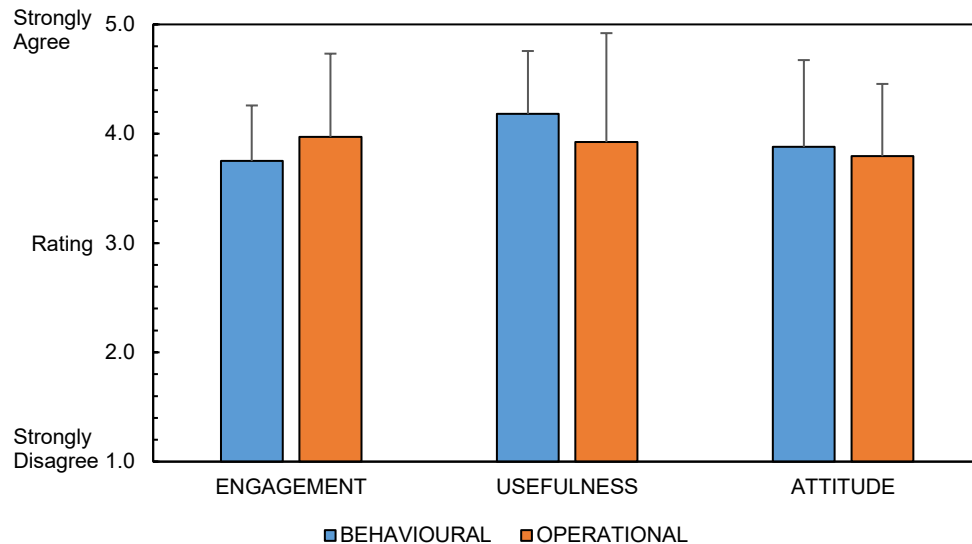


Figure 5-5 Mean aggregated TEQ ratings, scaled to match original scale values. Shown for for engagement, usefulness and attitude ratings with standard deviation error bars.

5.4.1.2 Trust

The total trust in automation questionnaire (TTAQ) (Gold et al., 2015b) comprises five sub-scales relating to: the discharge of the driver due to automation, safety gains, safety hazards, trust in automation, and intention to use. Again, ratings were captured using 5-point Likert scales, where 1 indicates strongly disagree, and 5 indicates strongly agree. Total trust was calculated as the cumulative score (i.e. the summation of all subscales). Participants were asked to complete the TTAQ before training and then immediately after driving.

Pre-training ratings were comparable between groups (Behavioural and Operational) for total trust and all subscales. However, by comparing ratings made before and after driving, it is evident that there was a significant increase in total trust (TTAQ) within the Operational group ($p = .02$) (Figure 5-6). Upon closer inspection, drivers in the Operational group indicated a greater intention to use automation after the experience ($p = .03$) and higher trust in automation (subscale) ($p = .003$).

There were no significant differences between pre and post ratings made by drivers in the Behavioural group (Figure 5-7, Figure 5-8).

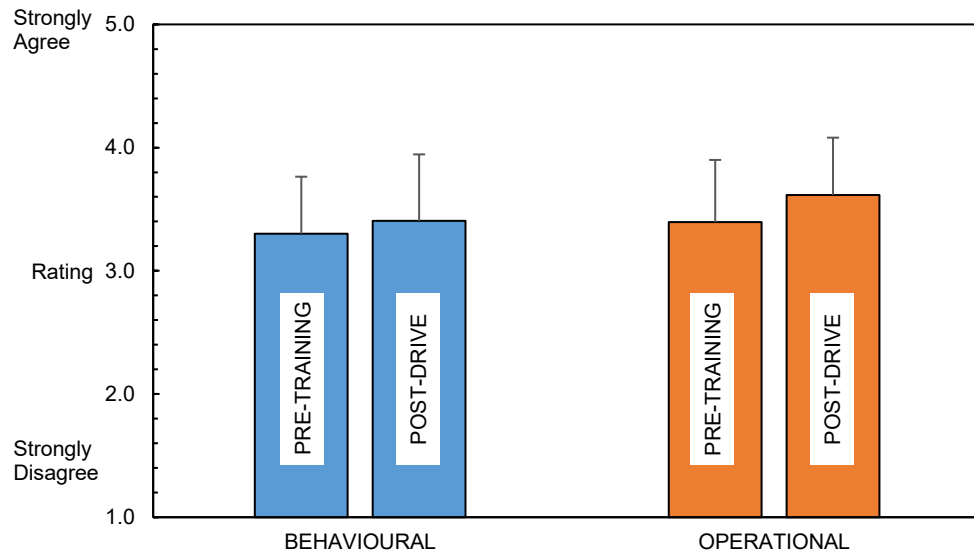


Figure 5-6 Total Trust (TTAQ) ratings mean values with standard deviation error bars.

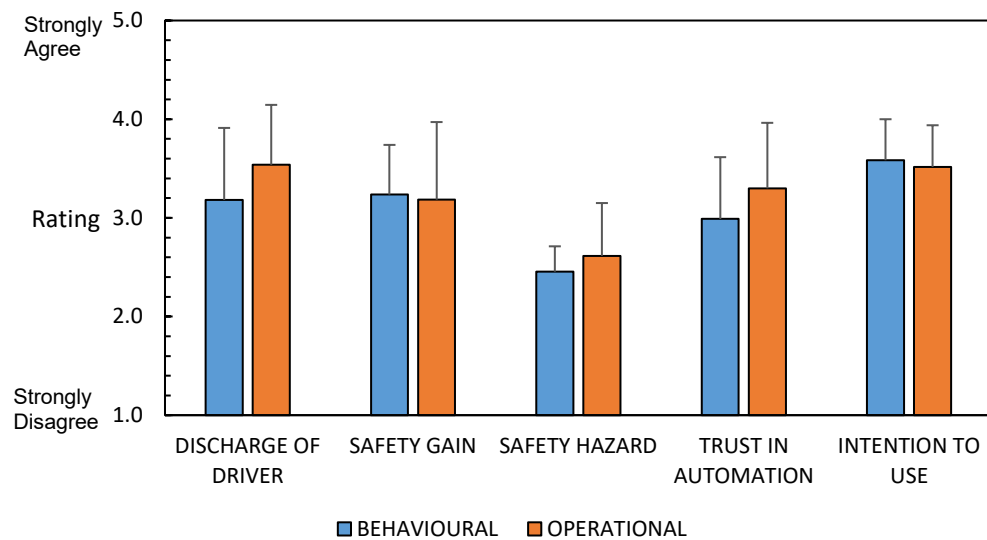


Figure 5-7 Pre-training trust ratings showing all subscales. Mean values with standard deviation error bars.

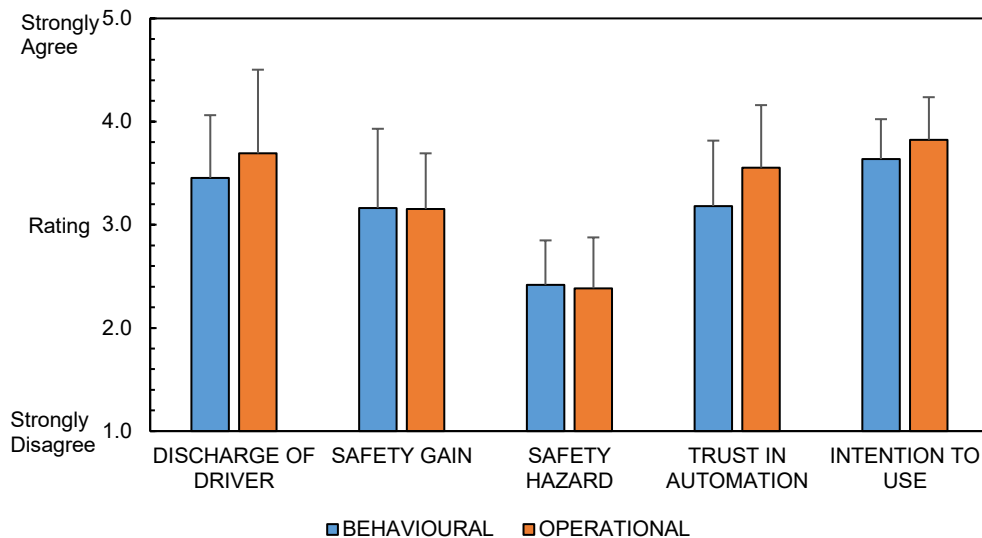


Figure 5-8 Post-drive trust ratings showing all subscales. Mean values with standard deviation error bars.

5.4.1.3 Situation awareness

The Situational Awareness Rating Scale (SART) (Taylor, 2017) explores respondents' perception of the attentional demand, attentional supply, and their understanding of the situation. Ratings are made using 7-point Likert scales, captured post-drive, where 1 is labelled 'low', and 5 labelled 'high'. Although ratings made by drivers in the Behavioural group were generally higher based on the responses captured during the study, the differences between groups were not statistically significant ($p = .097$) (Figure 5-9). There were no significant differences between groups for any of the subscales (Figure 5-10).

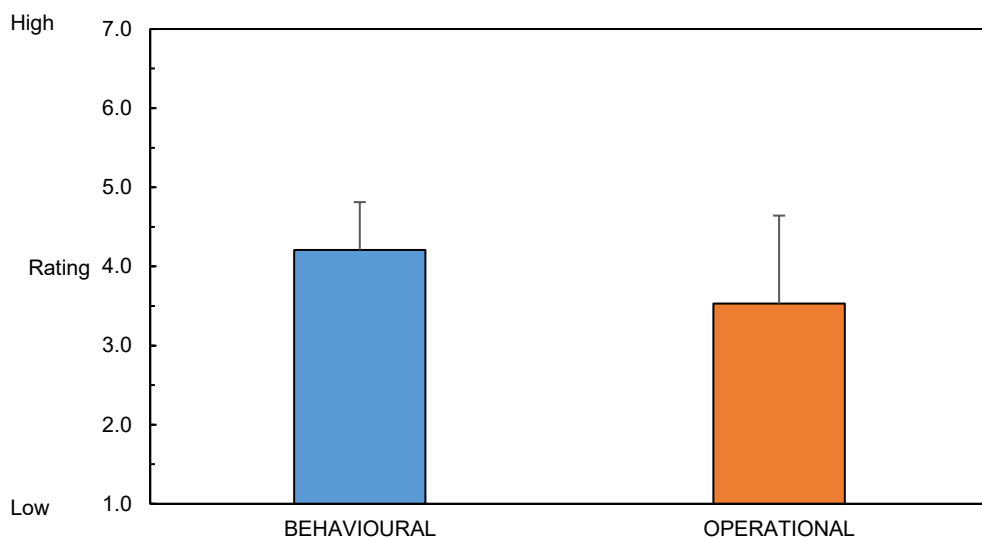


Figure 5-9 Aggregated Situation Awareness (SART) ratings, scaled to match original rating scores. Mean values with standard deviation error bars.

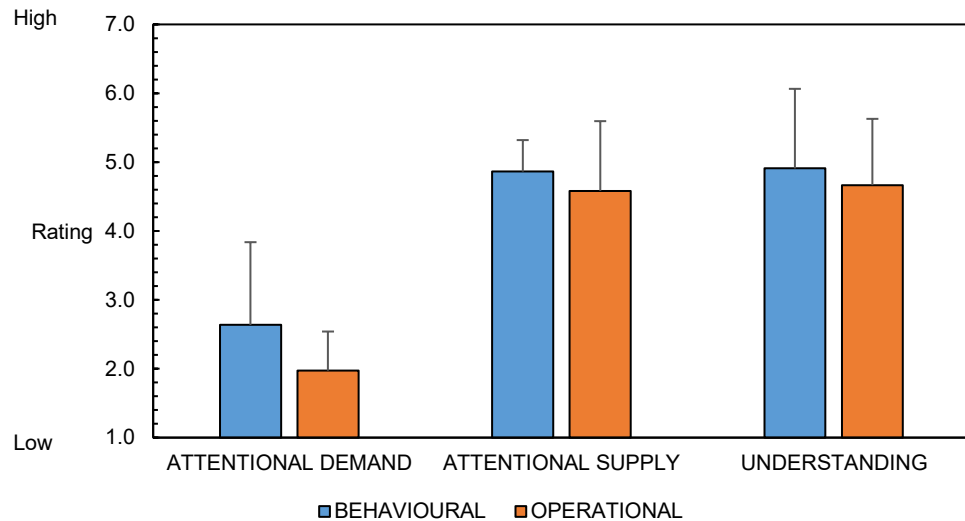


Figure 5-10 Situational Awareness (SA) ratings showing all subscales. Mean values with standard deviation error bars.

5.4.1.4 Workload

The NASA-TLX workload index (Hart and Staveland, 1988) is a multi-dimensional scale exploring mental demand, physical demand, temporal demand, effort, performance and frustration levels. Ratings were made using 7-point Likert scales, where a rating of 1 represents low workload, and seven, high. Although Total Workload (the numerical summation of all subscales, with ratings for performance reverse-scored) was statistically comparable between groups (Figure 5-11), those receiving Behavioural training indicated significantly higher temporal demand ($p = .048$) (Figure 5-12), suggesting they felt greater time pressure due to the pace at which the tasks or task elements occurred, compared to drivers in the Operational group.

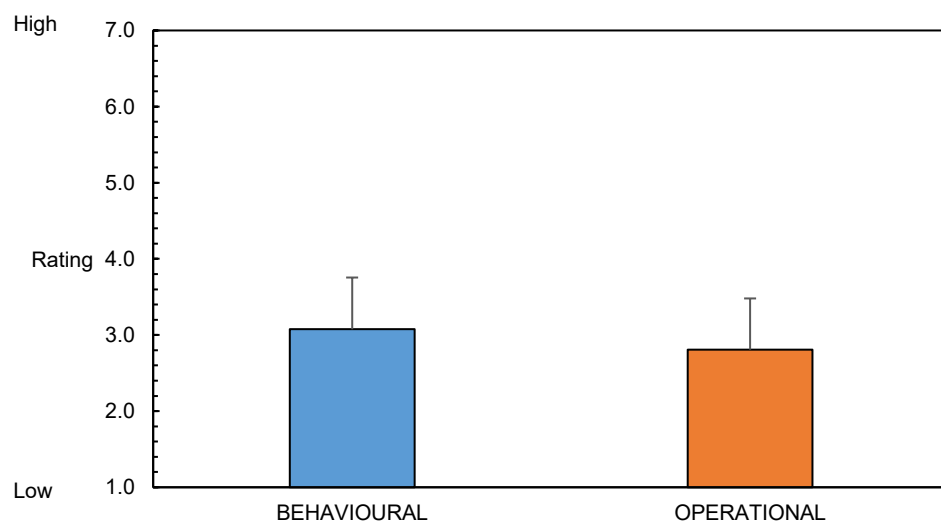


Figure 5-11 Total Workload (NASA-TLX) ratings. Aggregated ratings scaled to match original scale values. Mean values with standard deviation error bars.

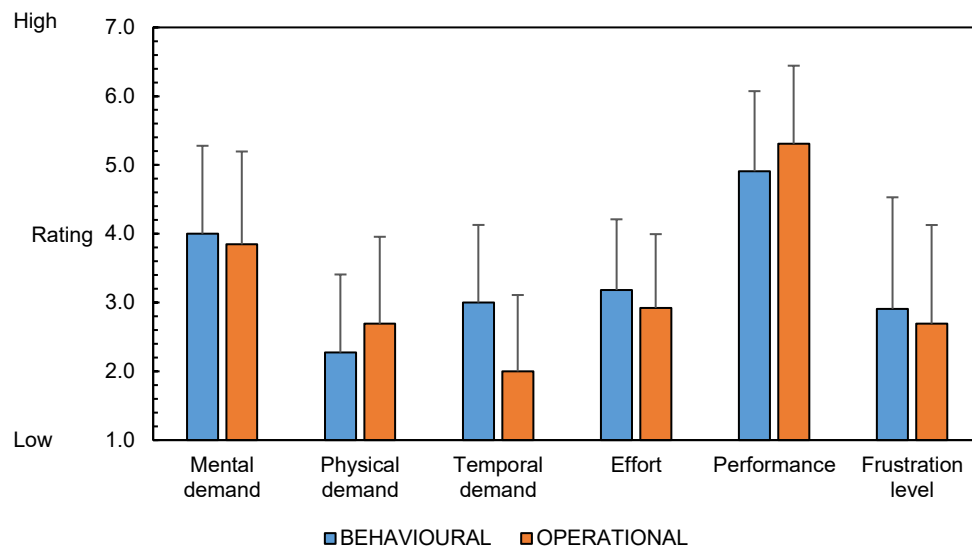


Figure 5-12 Mean workload ratings showing all subscales. Error bars show standard deviations

5.4.2 Visual behaviour

A key aim of the study was to change drivers' behaviour, not just their opinions. Therefore, it was expected that there would be a significant difference between cohorts (Behavioural versus Operational) with regards to their visual behaviour (e.g. number and frequency of glances at mirrors) during automated driving and following the resumption of manual control. In particular, it was anticipated that the Behavioural training cohort would demonstrate situation monitoring behaviours during the transition period that would suggest they were in an 'on the loop' state prior to taking back physical control of the vehicle, and this would have a knock effect on their situation monitoring behaviours following transition to manual control.

5.4.2.1 Glances to mirrors

The number of mirror checks were determined for all participants during the entire drive, i.e. during automated driving, when preparing to take over control, during the transition of control and after resuming manual driving. These provide an indication of the extent to which drivers attempted to remain engaged with the driving task, rebuild their awareness when asked to resume control, and the caution they applied during manual driving. The data show that drivers who received Behavioural training carried out significantly more mirror checks than those receiving Operational training during both automated driving and the transition to manual driving ($p = .010$ and $p < .005$, respectively) (Figure 5-13 and Figure 5-14, respectively). On average, drivers in the Behavioural group made 37.8 mirror checks during the 10-minute automated driving and 9.2 during the transition period, compared to 15.8 and 1.0, respectively, by drivers in the Operational group. In addition, drivers were statistically more likely to make at least one

glance during the transition of control in the Behavioural group ($p = .005$).

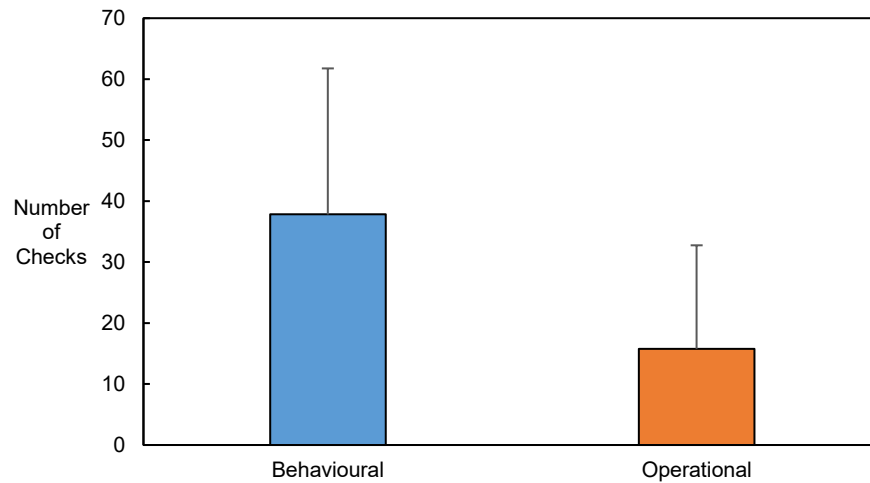


Figure 5-13 Mean number of mirror checks made during automated driving. Includes glances to left, right and rear-view mirrors. Error bars show standard deviations

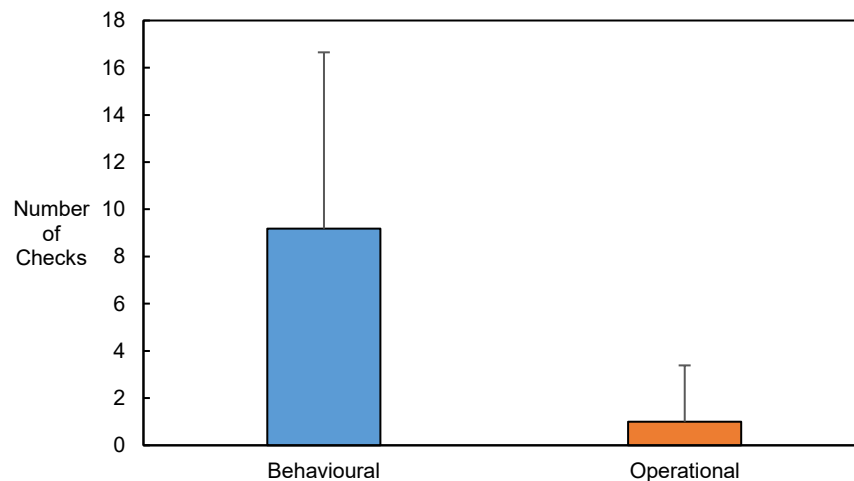


Figure 5-14 Mean number of mirror checks from prepare-to-drive notification (PTD) to manual driving. Includes glances to left, right and rear-view mirrors. Error bars show standard deviations

5.4.2.2 Awareness of the hazard car

As part of the experimental design, a car joined lane two behind the participant's vehicle simultaneously to the 'prepare-to-drive' TOR. The car sped up and remained close to the rear of the participant's vehicle, effectively tailgating it. The behaviour of the car was designed to replicate that of an irate driver wanting to overtake, i.e. abnormal behaviour that might be expected to draw attention and would normally be identified as dangerous or hazardous. The hazard car was plainly visible in the rear-view and side mirrors during the entire prepare-to-drive period (60 s). At the point drivers resumed manual control, the

hazard vehicle slowed and effectively disappeared from immediate sight. The intention in introducing the vehicle was to present another potential hazard to drivers during the takeover, but specifically one that required the use of their mirrors. Analysis of participants' glance behaviour shows that significantly more drivers in the Behavioural group (10 out of 11 drivers) saw the hazard vehicle ($p = .002$), compared to drivers in the Operational group (only 3 of the 13 drivers) (Figure 5-15).

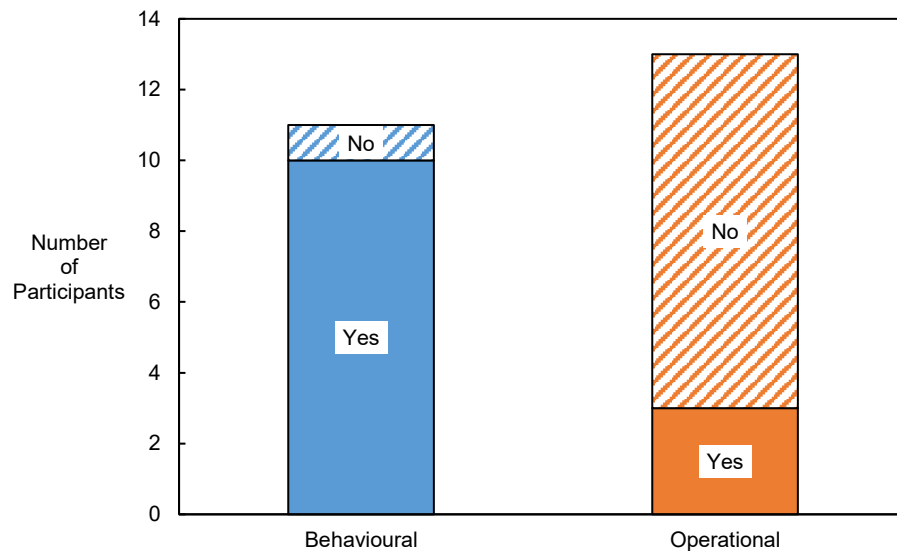


Figure 5-15 Number of participants who saw the tailgating, hazard car during the prepare-to-drive (PTD) period. 'Yes' indicates that participant saw the car, and vice versa

5.4.2.3 Shared attention

Although mirror checks suggest that drivers are actively attempting to remain engaged with, or to re-establish their engagement with, the driving scene, the way these were distributed, particularly during the 60-second transition/hand-over period, shows the relative importance with which drivers considered the driving task and their NDRTs. In other words, a glance to a mirror immediately after a takeover request (TOR) has been issued should indeed be welcomed, as it suggests that the driver recognises the need to see what's around them in preparation to resume control. However, if they immediately revert their attention back to their NDRT after this preliminary mirror check, it suggests that they have not fully disengaged with their NDRT and are subsequently not fully re-engaged with the driving task. Instead, such behaviour suggests these drivers are choosing to share their attention between both activities. In practice, drivers must discharge their NDRT as well as re-engage with driving before taking over control.

To explore this further, analysis was undertaken on the timing of the first driving-related glance (i.e. to a mirror or the forward road scene) immediately after the TOR was issued; any subsequent glances that

were associated with their NDRT; and, in particular, the final NDRT-related glance. Taken together, this revealed a period of ‘shared attention’ i.e. from drivers’ first attempt to re-engage with the driving scene (e.g. attempting to rebuild their situation awareness) to their last interaction with their NDRT. The data showed that drivers who underwent the CHAT behavioural training spent significantly less time sharing their attention during the prepare-to-drive notification period than drivers who underwent Operational training ($p < .005$) (Figure 5-16). On average, the former participants spent 1.8 seconds between their first glance to the road and their final engagement with or glance at their NDRT, compared to 11.2 seconds for the latter. Moreover, the time to the first mirror check following the PTD was also significantly different between groups ($p < .005$), with Behavioural participants taking on average 7.3 seconds before making their first glance at the road, whereas Operational participants took 21.3 seconds. Even so, it is worth highlighting that some drivers – in both groups – were not engaged with their NDRT at the time the TOR was issued. Five of these were in the Behavioural group, and three in the Operational group.

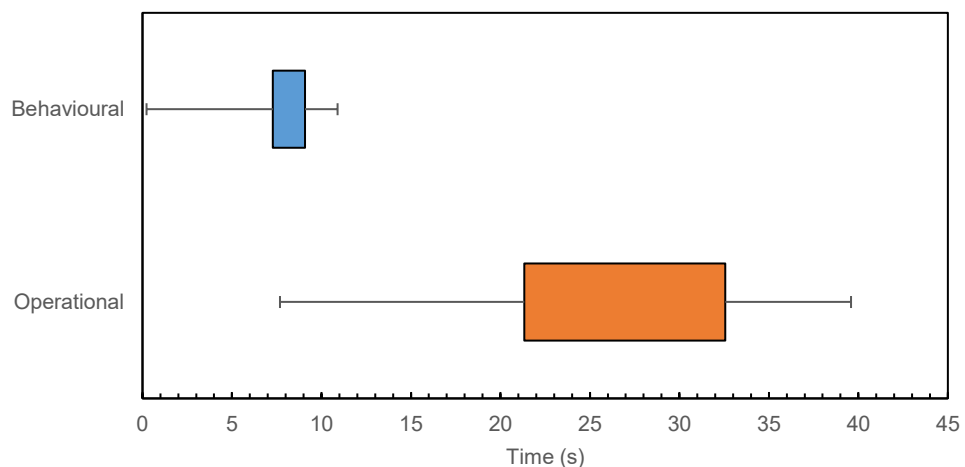


Figure 5-16 Initiation and duration of shared attention during prepare-to-drive (PTD) showing first driving-related glance (left end of solid bar) to final NDRT-related glance (right end of bar). Mean values with standard deviation error bars

5.4.2.4 Visual behaviour after resuming manual driving

Drivers’ visual behaviour remains important after resuming manual control, not least due to the fact that, in the current study, they were required to move into lane one – meaning they needed to negotiate with other road users already present in lane one – and then exit the dual carriageway. All mirror checks were subsequently identified from the resumption of manual driving. These were categorised as either glances associated with rebuilding situation awareness (SA), or were specifically attributed to the lane change manoeuvre (LC). The latter were defined as those which immediately preceded physical actions

associated with manoeuvring their vehicle, based on our expert evaluation of the videos. These actions included repositioning hands on the steering wheel (for example, to activate the indicator), or beginning to turn the steering wheel to guide the vehicle into lane one. From these actions, a lane-change timeline was identified, and this was subsequently cross-referenced with the vehicle performance data captured by the STISIM simulation software.

It is noteworthy that, on average, the time from resuming manual control to the first LC mirror check (i.e. immediately preceding the commencement of the LC manoeuvre) was the same (i.e. there was no statistical difference) for drivers in both groups ($p = .74$). However, the delay before the lane change was physically enacted following the final confirmatory LC mirror check (effectively, the time taken to begin to make the LC manoeuvre after deciding to do so) differed between groups ($p = .02$), with drivers in the Behavioural group taking significantly longer than Operational drivers (mean times: 4.3 seconds and 2.2 seconds, respectively), suggesting greater caution amongst the former drivers. Nevertheless, there was no significant difference in the time taken to physically manoeuvre the vehicle from lane two to lane one once this action was started ($p = .36$) (Figure 5-17). In addition, drivers who received Behavioural training made significantly more mirror checks prior to and during the lane change manoeuvre ($p < .005$ and $p = .012$). These drivers were also significantly more likely to make multiple glances to their mirrors during the lane change manoeuvre itself ($p = .016$).

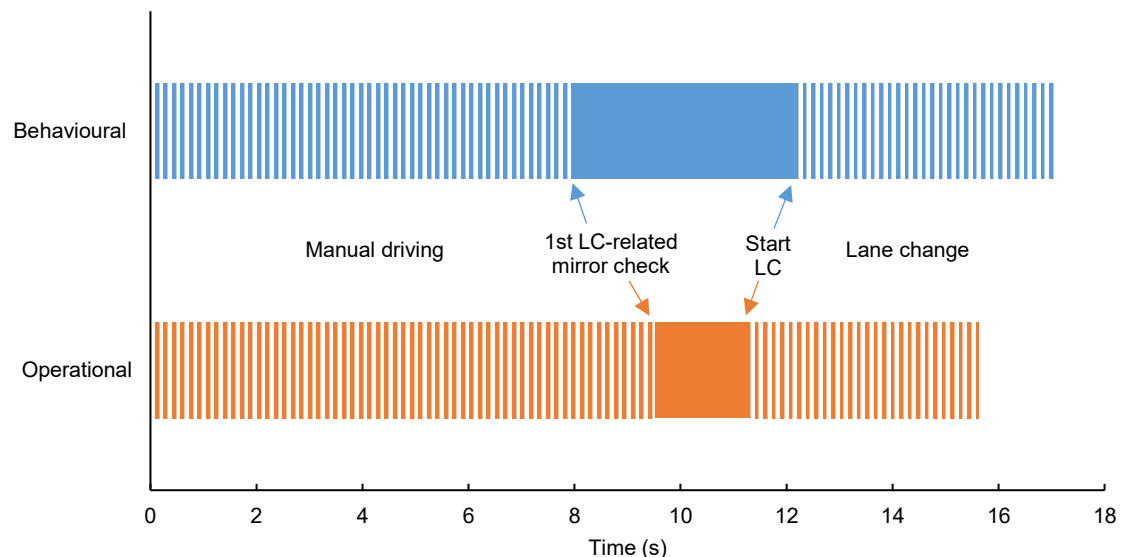


Figure 5-17 Time to contemplate and undertake lane-change (LC) manoeuvre (mean values). Time zero represents the start of manual driving

5.4.3 Driving performance

A key factor in any study investigating the transfer of control from automated to manual driving is drivers' manual driving performance immediately following the takeover. In the longitudinal L3 ADS driving simulator study by Burnett et al. (2019), lateral control (lane swerving, in particular) was shown to be poor during the 10 seconds immediately after resuming control. However, as reported in chapter three, the findings of the observation study on the data from this longitudinal study, showed that this control notably improved following repeated, daily exposure across the five drives. During the current study, the focus was on changing drivers' behaviour by improving their knowledge and awareness (for example, by completing and/or correcting their mental model). The expectation was that these drivers will subsequently act more appropriately during automation and when asked to resume control. However, it was also anticipated that this should improve their driving behaviour and performance immediately following the transfer of control. Therefore, driving performance data was captured from the STISIM simulation software, and key driving performance metrics were used to explore the rate of change of longitudinal and lateral behaviour, i.e. longitudinal and lateral acceleration, standard deviation of lane position and steering reversal rate, aiming to identify the smoothness of the resumption of control and lane change manoeuvre. In addition, the minimum time to collision (TTC) was calculated to provide an indication of how close participants were to vehicles in their immediate surroundings.

Longitudinal and lateral acceleration were calculated in 1-second time intervals following the resumption of manual driving control. The data suggested no apparent difference between the Behavioural and Operational groups ($p = .48$ and $.24$, respectively, when averaged across the entire 10 seconds). Comparing each 1-second interval between groups, there was some evidence of statistical differences at $t=2s$ and $t=6s$ ($p = .048$ and $p = .011$, respectively), with data suggesting that drivers in the Operational group demonstrated higher longitudinal and lateral variability (respectively). However, in isolation, these two differences were largely inconclusive. Additionally, there was no evidence of any statistical differences between groups for standard deviation of lane position (SDLP) over the 10 seconds following the takeover of control ($p = .56$), or the minimum time to collision (effectively, the distance between the participant's car and other vehicles on the road) ($p = .95$). Steering reversal rate was also comparable between groups ($p = .90$).

5.4.4 Non-Driving related tasks (NDRT)

As with the longitudinal study by Burnett et al. (2019), drivers were able to select and bring with them their own non-driving related tasks and activities (NDRTs). Drivers were given the freedom to choose how and

when to engage during automation and no restrictions were applied to the type of tasks and activities that drivers could undertake. Whilst this approach arguably fails to provide standardisation from which to compare driver takeover performance, it nevertheless increases the ecological validity of the experience and participants' behaviour, not least because it is expected to evoke important motivational factors that could have a significant bearing on drivers' engagement (or lack of) with the driving task. Although the type and range of different NDRTs were not formally analysed as part of this study, general observations made that suggested drivers used a range of NDRTs. The majority of drivers, predominantly used mobile phones, tablet computers and laptops, but a number also brought books and printed papers with them to read during automation.

5.4.5 Summary of Questionnaire, Behavioural and Performance Results

Table 5-2 Summary of results from subjective measures

Category	Measure	Result
Training Evaluation Questionnaire (TEQ)	Engagement Usefulness Attitude	No significant differences between Behavioural and Operational training
Total Trust in Automation Questionnaire (TTAQ)	Total Trust (TTAQ)	No significant differences between groups pre-training Ratings significantly higher post drive for Operational group ($p = .02$)
	Trust in Automation Intention to Use Discharge of the Driver Safety Gains Safety Hazards	Significant increase in Trust in Automation for Operational group ($p = .003$) Significant increase in Intention to Use for Operational group ($p = .03$) No significant differences in Discharge of the Driver, Safety Gains or Safety Hazards
Situation Awareness (SART)	Attentional Demand Attentional Supply Understanding Situation	No significant differences between groups
Workload (NASA-TLX)	Mental demand Physical demand Temporal demand Effort Performance Frustration level	Higher Temporal Demand for Behavioural group ($p = .048$)

Table 5-3 Summary of results from objective measures

Category	Measure	Result
Visual behaviour	Glances to mirrors/roadway during automated driving and prepare-to-drive (PTD)	Significantly more mirror checks during automated driving and during PTD by Behavioural group ($p = .010$ and $p < .005$). Statistically more drivers in Behavioural group made at least one mirror glance during transition of control ($p = .005$) Significantly shorter time to first glance at mirror/roadway during PTD ($p < .005$) for Behavioural group Behavioural group statistically more likely to see the tailgating blue car ($p = .002$)
	Shared attention during PTD (driving / NDRT)	Behavioural group spent significantly less time sharing their attention between driving and NDRT during PTD ($p < .005$)
	Glances to mirrors/roadway during manual driving and lane change manoeuvre (LCM)	Behavioural group undertook significantly more mirror checks before starting LCM ($p < .005$) Behavioural group undertook significantly more mirror checks during LCM ($p = .012$) Behavioural group significantly more likely to make multiple mirror checks during LCM itself ($p = .016$)
Driving behaviour	Lane change manoeuvre	Behavioural group took significantly longer to begin lane change following the final confirmatory mirror check ($p = .020$) No significant difference in the time taken to physically change lanes ($p = .36$).
Driving performance	Manual Driving	<i>No significant differences between groups identified for:</i> Lateral and longitudinal acceleration ($p = .48$ and $p = .24$, respectively) Standard deviation of lane position ($p = .56$) Steering reversal rate ($p = .90$) Minimum time to collision ($p = .95$)

5.4.6 Thematic Analysis of Semi-structured interview

An inductive thematic analysis (Braun and Clark, 2006) was undertaken on the transcribed interviews to identify themes from the data that related to the thesis research questions. The themes are influenced in

part by the probe questions that consistently drew out key information from drivers in both training conditions:

- What were you seeing, hearing, noticing?
- What information did you use in making this decision or judgement?
- What let you know that this was the right thing to do at this point?
- What training might have offered an advantage in this situation?
- If a novice had been in charge at this point, what type of error might they have made and why?
- What knowledge, information, or tools/technologies could have helped?

There was a focus on drawing out evidence to support the underlying proposition that there is a necessity for behavioural training or guidance to support drivers in transitions from automated to manual driving.

Additionally, themes that focused on identifying the cues and information drivers drew on to shape their behaviour, driver mental models and expectations surrounding their responsibilities during the transition of control, and driver views and ideas on the tools, technologies or training that could be advantageous to support resumption of the driving task, were drawn from the data. Commentary on each of the themes is provided below, along with analysis contrasting the views of drivers between the behavioural (BT) and operational training (OT) groups and relating to relevant points within the literature:

5.4.6.1 Engaging automation

Drivers were asked what cues and/or information they used to decide to engage the ADS. A common theme reported by drivers was the importance of checking their surroundings before engaging automation. Drivers reported attending to the driving environment and using that information to guide whether it was safe to engage the ADS:

“it was safe and it was clear and that told me it was okay ... to engage automation ... if it had been busier ... I may have waited for the road to quieten down.” [P18, OT]

“I was aware of my surroundings ... I needed to see if it was safe to ... get into the automation mode and that is only possible if you know your surroundings especially ... what’s happening behind.” [P11, OT]

“[it was a] straight road, nothing was behind, it was safe to engage ...” [P17, BT]

“I was looking behind and ... in front of me ... there was a clear road ... it is safe ... this is the right time to engage” [P14, BT]

Interestingly, regardless of the type of training they received, drivers more readily and consistently recalled using their SA of surrounding

vehicles, especially vehicles to their rear, to inform their decision to move from manual to automated driving than when they were asked the same question about transitioning from automated to manual driving

5.4.6.2 Information and Actions for Driver Readiness

Drivers were asked various questions to probe the reasoning behind their actions following the 'Prepare to Drive' notification on the in-vehicle HMI. For example, drivers were asked what they were hearing, seeing and noticing at the point the TOR was issued and what information drivers used to decide that they were ready to drive.

There was a difference in the information and cues drivers were focused on depending on whether drivers had the CHAT (behavioural) or operational training.

Drivers in the CHAT group discussed actions to prepare themselves physically within the driving cab and also emphasised a focus on cues from the external environment. They highlighted the need to make checks to the driving environment with a view to raising their SA in relation to surrounding vehicles and to begin planning future actions:

"First I had to ... be in control of my body so I was ... moving [my body] ... feeling that I was ready to engage again ... approach ... the steering wheel so more of a physical contact as well so I was extending my legs again ... and ... I had to be ready mentally"
[P14, BT]

"I looked in all the mirrors... I was trying to see what was happening around me ... you can see in front of you that's your main, but ... when you're changing speed behind you is ... more important because you don't know exactly what's going on behind you ... I knew I was going to have to take a left. I might have to speed up passed someone and jump in or slow down and jump behind." [P22, BT]

"I was thinking about different manoeuvres ... I was observing the car behind me ... the cars on the left and then making the decision ... if I'm here ... I can pull to the left lane ... or not [and] overtake another car" [P23, BT]

Whereas drivers in the operational group tended to focus on the HMI and, in particular, the timer:

"the ... display ... was the obvious one ... and then ... there was a voice saying you can take over." [P13, OT]

"I'm seeing the counter coming down and then hear the beeping sound of the 3, 2, 1." [P3, OT]

"I was conscious of how much time had elapsed on the timer, I could ... see the bar coming across so I knew I had plenty of time and then I think I was just kind of back into normal driving mode"
[P19, OT]

Drivers in the operational group acknowledged that they could have done more to prepare themselves but did not express a need or an importance to regain SA of the driving environment like the CHAT group did. Their focus was more on regaining physical control of the vehicle. Additionally, when these drivers did conduct checks, they tended to be to the front of the vehicle and mainly with a view to controlling the speed or the steering during the transition:

"[I was] Noticing basically just to see when the manual drive mode is actually engaged, when do I actually get the controls... [anything else] no ... I could have looked for hazards to see ... why it was telling me to take control." [P3, OT]

"I just used the information that was on the screen ... I could have assessed the situation outside myself as well ... I did not pay much attention to what the surroundings were ... I wasn't doing anything so I could ... take control." [P2, OT]

"There was nothing in front of me and ... the screen with the countdown was telling me that I was soon to be in control and just sort of the positioning of things such as the steering wheel ... to make sure that when I did take control ... it would be a smooth transition." [P21, OT]

There was a temporal difference between the way drivers in the CHAT group spoke about their actions during the transition period that suggested a recognition of the need for repeated checks and an extended time conducting 'on the loop' behaviours:

"It wasn't good enough just to assume there was nothing new because I'd been monitoring the situation I wanted to double check that there was nothing new that I needed to attend to." [P20, BT]

"I had to be ready mentally ... check ... again the cars on my left ... again in front of me, there was no one so again a safe situation and then check behind." [P14, BT]

Whereas drivers in the operational group spoke about checks in a singular, transient and cursory way:

"I think just ... knowing your speed a very quick look up, there were no cars anywhere near me, centre of the road so I'm ready." [P19, OT]

"My phone's away and my hands are back on the wheel kind of been looking out for a few seconds kind of feel I know where I am." [P20, OT]

Some drivers from the operational group highlighted that the length of the transition time gave them comfort and time to prepare:

"Having that prolonged period before starting to drive was really useful ... because I knew it said prepare to drive and I was like okay I've got a minute until I need to drive ... I ... can get my bookmark out and ... get rid of my book ... then I can deal with my coffee and then I can get with driving ... my first instinct was ... how am I going to do this I've got a hot coffee here you know aaaahhh and then it's like no you've got some time ... just get yourself sorted." [P21, OT]

"I knew that I had a bit more time ... I could see the bar moving across the screen ... I knew that there would be a ... count down at the end ... I felt more comfortable ... getting prepared for it." [P1, OT]

However, some drivers in this group continued using their NDRTs after the TOR and subsequently described experiencing temporal load when manual mode engaged and the potential for startle responses or as reason for not making checks:

"The moment I take over ... all my senses will be active and I'll have to be in some sort of stress to take control of the car and watch everything going on." [P11, OT]

"If you've been relaxed and engrossed in other activities ... and then ... suddenly it's saying you're now in manual takeover ... would your reaction times do the stuff that you would do automatically, like ... checking in your mirrors and braking or do you just ... slam your foot down on the brake without even looking behind you." [P5, OT]

Drivers were given the option to override the minute long transition period by pressing a button that would immediately move to a 10 second count down to manual driving mode engagement. Drivers who pressed the button were asked what information they used to make that decision. Drivers in both groups took this option but they provided different rationale for having done so or having taken different actions before pressing the button. Drivers in the CHAT group tended to conduct checks to the external environment before considering whether to move to the 10 second countdown with some linking it to a challenge of maintaining awareness without being in the motor control loop:

"Once you've got all the awareness and ... you've done all the checks there's no point in delaying because ... you're ... having to

maintain that level of awareness without the control which is ... harder to do than ... once you've taken control and so I decided that I should take control at that point" [P20, BT]

While others linked it to decision making either regarding hazards in the current environment or future actions:

"I first had a full look around so I was checking all the mirrors and behind and I could see behind there was a car fast approaching in terms of speed, so I was just waiting a bit more of distance before I engaged because I didn't know ... how fast I could manage the speed and then there was a car on the left, so I was just waiting to pass the car" [P14, BT]

"I ...pressed the button before the 50 second because ... I was checking ...I knew that behind I was okay, on my left there was not such a density of traffic in front of me was fine and I was already aware that there was a road to exit so ... it all makes sense ... I've got all the information I can take over I know what I need to do." [P7, BT]

Whereas the operational group gave reasons that tended to be nonchalant and focused on gaining control faster, but without a clear rationale or plan:

"I had the option to press the button and engage it quicker than for me to wait ... so that's why I pressed it." [P3, OT]

"I was a little ... bored so I thought I'd ... engage myself as soon as possible." [P13, OT]

"Why wait, there's no reason for me to wait I'm prepared so I can go" [P25, OT]

"I ... heard what the car said ... I just pressed [the button] ... I wanted to take control straight away ... the voice raised alertness in myself." [P19, OT]

A few drivers in the Operational group stated that they believed they had conducted checks during the prepare to drive transition period. However, as the recording of the drive was used to aid the interview, it was possible to probe the driver in a way that highlighted this recall error and explore the driver's reflections on this scenario (see **Error! Reference source not found.**):

Int: What actions had you taken to prepare yourself to drive at this point?

Driver: I had put my phone and papers aside, got my foot on the pedal and ready just to go and hand on the steering wheel

Int: Okay anything else?

Driver: Um I think I looked around as well ...

Int: okay let's just rewind it back ... so you got prepare to drive there ...

Driver: I'm thinking there might be some ... vehicles as well behind me

Int: were there?

Driver: I'm not sure I was just thinking that [while watching the playback]

Int: did you see any?

Driver: No I didn't see anything no

Int: Okay did you see the vehicle behind you?

Driver: the one I just drove passed

Int: No the one behind you in your lane ...

Driver: Oh! I didn't see it!

Figure 5-18 example conversation between interviewer and driver [P25, OT]

Drivers were surprised that they had not conducted thorough checks to the driving environment and also reflected on what they would do differently:

"I would have let the timer run down further ... and I'd be more alert." [P9, OT]

"While you're waiting [for manual mode] ... you tend to sort of look ... instead of concentrating on the road ... I tend to look ... at the screen [HMI] ... I didn't look in my mirror either to see if anything was behind me I was focusing on that one in front." [P5, OT]

5.4.6.3 What training would be advantageous?

There were varying views from the drivers regarding the need for training in vehicles with L3 ADS. However, this question was posed a few times at different points during the review of the drive and phrased both in terms of the views of the driver's own needs and in 'what if' scenarios e.g. what if it was a novice driver. Depending on this timing and phrasing of the question (i.e. in relation to novice drivers), the drivers gave different answers as well as having a different viewpoint depending on the training they had received for the study.

Drivers in the operational group tended to feel that there was no need for training and experience would be enough:

"I don't think that I would need any training ... if it was me, I would probably just learn by using the features ... and once I am familiar ... I would be confident in using them more often." [P3, OT]

"I think it would just be case of getting used to something you would probably get more relaxed ... leave it to the last minute ... knowing exactly what time you've got, so I don't know whether training would make much difference." [P1, OT]

Some based this view on the belief that they only needed to take over the operational tasks of the driving, which they viewed as simple to do:

"Additional training? No, I guess it's just I think you need like important point in the engage and disengaging and I think that you know once it engages or disengages you can tend to swerve a little bit or accelerate or decelerate." [P19, OT]

However other drivers in the operational group expressed that they felt training would be important, in particular for the transition back to manual control and that drivers might not realise the importance of certain actions and make errors through not conducting checks:

"You'd ... go back to the basics, I think people can get ... complacent with their driving um so I think you'd have to ... go back and you know check your mirrors, check your blind spot, signal ... that sort of stuff" [P13, OT]

"I think the training should be given ... it's not about driving the car ... but the transition from manual to automatic and automated to manual I think that requires ... some sort of training... for change from manual to automated ... this could be trickier actually because ... you need to be careful about your surrounding and ... I believe people will make more mistakes here" [P11, OT]

"Some instruction into the best way of doing that transition ... have a good look around get yourself centred and know where you are again before you do any manoeuvre ... you should be doing that during the countdown ... you've got to get yourself in the car again before you're expected to be in control again" [P4, OT]

Some drivers in the CHAT group felt that training should be incorporated into the curriculum for learning to drive. They highlighted the potential of high workload and cognitive tunnelling leading to errors such as inattention blindness:

"It would definitely have to be incorporated in your driving lessons ...about taking back control of the vehicle ... it's unfamiliar ... people would be worrying too much about the speed and stuff and then they won't take it in the round ... they focus on the one thing where actually that's ... not the whole picture ... so you could say I'll maintain the same speed, but they're worrying too much about that and then they don't notice that there can be a motorcyclist coming up ... they won't naturally look round..." [P24, BT]

Within both groups, there were drivers who changed their views on “what training would be advantageous?” after reviewing their experimental drive. This change appeared to be triggered when these drivers realised that they had missed key information due to not undertaking checks during the transition period (see **Error! Reference source not found.** and **Error! Reference source not found.**):

Before review:

“I don’t think there is any training needed for that ... maybe just letting them know how to position their hands and to be ready to press on the accelerator, brake just ... emphasising more the ... you have to physically steer the car after”

After review:

“I think in that [transition] period it’s pretty important to look around and familiarise yourself with the cars ahead, behind and all your mirrors.”

Figure 5-19 Example 1 of driver comments regarding training before and after the review of the video of their transition of control in the simulated drive [P9, OT].

Before review of TOR and transition:

“I think that depends on the driver ... for someone like me who drives a lot ... I think reading the manual was good enough. I think for someone who ... doesn’t drive as much I think ... more training or ... some ... accompanied drive ... would be helpful just to give them confidence ... if it’s someone who’s experienced ... I wouldn’t make it compulsory ... but for others maybe.”

After reviewing the TOR and transition when the driver realised they hadn’t checked their mirrors:

“Check your mirrors [laughing] possibly depends on what training would be offered ... probably a reminder of what you should do before taking control of the car again ... some kind of information or alert about the changeover process and how if you’re not fully prepared ... you need to be”

Figure 5-20 Example 2 of driver comments regarding training before and after the review of the video of their transition of control in the simulated drive [P21, OT].

One driver had strong views on the CHAT training prior to reviewing the takeover:

“I thought the training was a bit like teaching your grandmother to suck eggs ... checking everything is what you do when you’re

driving all the time anyway so basically it was telling you what to do when you're driving all the time... it was deconstructing a skill."
[P20, BT]

However, when the researcher asked, *"what information was telling you, okay these are the actions that I need to take in order to get myself ready?"* the driver responded, *"It was what I had remembered from the [CHAT] training."* (P20, BT) and when discussing potential errors novices might make during the transition from automation to manual driving their views on training evolved:

"Driving is a skill and as such ... you're not conscious of all the things you're doing so when you take over a car in the middle of a drive that's when you become aware of all the things that you've suddenly got to start doing when you haven't been doing [them] 2 seconds previously ... I think people could well be overconfident ... the first time they do those manoeuvres." [P20, BT]

There were various themes that arose relating to the style of training the drivers had undertaken for the study. Drivers who underwent the CHAT training felt it gave them something new to think about:

"The training gave me something to think about ... I think it was good to focus you on ... what you could do to drive safer ... the training made me ... more aware ... of what is expected of you."
[P24, BT]

"The training ... gave me that schema of there being a level 1, 2, 3, 4, 5 ... I was aware that it wasn't a level 4 vehicle so I should still pay attention and then the checks that it told me to do ... I thought I'll do those checks, so it was useful." [P10, BT]

Whereas drivers who underwent operational training had views on the utility of user manuals for training purposes. This tended to be that although the information contained in user manuals was useful, it was unlikely to be a sufficient, standalone form of training as most drivers didn't read them:

"The information [operational manual] I got before I started was really useful ... I think training physically like to drive more in this sort of car will provide more confidence, and trust." [P2, OT]

"User manual ... I'm not going to read it and I don't think that most people do either ... it would have to be a much more engaged level ... I think it'd have to be ... compulsory and making sure that people have done it, and done it, and done it again" [P15, OT]

"Manual documents are very hard to go over you're not used to reading them ... people never get around to actually doing it ... but information like yeah it would be good to have like good information." [P25, OT]

Concern was also raised about the use of mandatory, regulated training for experienced drivers:

"I can't imagine people going to a test centre or at the age of 40 going to their driving instructor just to learn that skill or process ... I'm not sure if there's anything ... official ... just repetition of ... going through it." [P13, OT]

The range of views from the drivers suggests that there may be a need for a tailored and/or layered approach to driver training. For example, different training approaches might be more appropriate for different groups, such as mandatory training for novices versus test drives or operational manuals for experienced drivers. However, a layered approach where training or guidance is delivered through a variety of means, may need to be taken to ensure training reached driver groups, especially more experienced drivers. As drivers are unlikely to read the operational manual and typical content user manual documents are not sufficient to guide appropriate driver behaviours to support safe and effective automated to manual transitions.

5.4.6.4 What if it was a novice driving?

When discussing 'what if' scenarios and, in particular, what errors novices might make, most drivers highlighted issues related to SA, increased workload and making errors due to startle effects or lack of information about the driving environment. It appeared that talking about the experiences of novices, as opposed to their own performance, made it easier for drivers to highlight the potential pitfalls that they were also susceptible to:

"Not matching the speed ... a jerky jitteriness on the steering wheel ... flusteredness of oh gosh I've got to drive again now ... perhaps not being fully aware of the surroundings and what might be behind you and the dangers and the risks involved ... perhaps even starting that manoeuvre before clearing the car that was ... next to me." [P21, OT]

"Failing to check the position of other vehicles ... pull in without checking anything on the left-hand side ... you're approaching a junction ... if you're a novice then oh I've got to get this exit I'll go and lack of checking in the mirrors before doing so" [P40, OT]

However, some drivers highlighted that issues or errors could be related to complacency rather than inexperience:

"I actually think it would be more of a problem to people who are more experienced ... cause you're ... used to driving in a particular way. I think if you've not learnt to drive at all and this was normal it'd become second nature you'd be very used to it

everything feels weird when you're learning to drive um and you kind of get set in your ways a bit I guess." [P12, BT]

"I think there's ... a strong chance of an incident by just not driving ... according to the traffic because you've gone from not being in charge ... to oh the car's going at 70 mph and all of a sudden I'm in charge ... I guess ... you don't really check your mirrors like you should do ... even when you're setting off at home ... after a while you don't." [P21, OT]

Using 'what if' scenarios was a useful way to highlight some of the challenges and potential areas training or guidance for automated to manual transitions could focus on. A key point of concern raised by both the CHAT and Operational groups was that drivers, regardless of their experience levels, may not carry out checks prior to receiving manual control and, therefore, make decisions with little SA while experiencing high workload. In their responses, drivers seemed to be aware of what they needed to do but also reflected that there was a good chance that they would not do it either through lack of experience or developed complacency with lots of experience.

5.4.6.5 HMI Feedback

There was a variety of views on what information the HMI could report to the driver that would be useful during the automated drive and in the transition to manual driving. A key point from many of the drivers was that, while the ADS was controlling the driving task, they wanted a balance between being given reassurance that the vehicle was in control, but that they did not want to be saturated with information:

"It would be nice to know ... that it's [the car] aware of everything ... but ... the whole appeal of automated driving is that I can just let the car get on and do its thing and I can get on ... so if every 5 seconds it's ... saying oh there's a car passing here, it's going to get really annoying and you're not going to be able to focus." [P21, OT]

Drivers suggested that the HMI could provide guiding principles to help them carry out actions to transition smoothly and safely from automated to manual mode. Some drivers felt that a reminder of required actions would support drivers due to potential complacency effects:

"I think it could probably remind you of the [CHAT] manoeuvres ... because ... being more comfortable with the car ... I'd probably ... not forget to do the manoeuvre but trust it so much that I don't really do any of the checks and then that could cause a problem ... that could get annoying but it's better to be safe than sorry I suppose." [P10, BT]

Views of drivers from the operational group included those who wanted real-time guidance provided to them, while they built experience with using the system to help them to focus on the takeover:

"It will guide you ... step by step ... I wish that there would be some screen where there are some steps ... when I'm using it for the first few times, it will make sure that I have full concentration."
[P3, OT]

And those who felt that guidance would help to support decision making with the inference being a recognition that there would be a spike in workload during the transition:

"More general information of what would help ... you to ... make the correct decision by having more information available to you rather than ... now I'm in charge and then you ... think what's in front of me, what's behind me and all these things are going on and you might have to make a split-second decision." [P5, OT]

The information drivers wanted from the HMI was influenced by their mental models of ADS and subsequent expectations of the driving task responsibilities during and after the transition to manual control. Drivers' mental models influenced both their views on the actions they should carry out following the 'prepare to drive' takeover request and the information they felt should be provided to them by the vehicle sensors and HMI.

Some drivers reported recognition that they would have a lower level of awareness of their surroundings due to engaging with NDRTs. However, they viewed the ADS as having continued responsibility for checking their surroundings, even though they had been issued a takeover request, suggesting a diminished level of responsibility for the driving task:

"I think ... if there is a car behind me there should be a ... display, some sort of alert ... although I'm giving you manual mode back still you have traffic behind because if you are involved in a certain activity you are not aware of your surrounding ... if the system can detect that there is an obstruction coming and you have to switch to manual it should also tell you what's in your immediate surroundings ..." [P11, OT]

There was some recognition of the requirement to increase their awareness of the driving environment following the takeover request and that there was sufficient time for them to carry out checks during the transition period. However, some drivers still felt these checks should be carried out by the ADS and communicated to the human driver. Incidentally, it was drivers who had not carried out their own checks during the prepare to drive transition period who reported these views after reviewing the experimental drive:

"The system requires me now to watch all the mirrors and see what's happening before I grab the steering [wheel], although there was sufficient time, but ... if there was ... assistance that can just tell me about my surrounding then that would be advantage."
[P9, OT]

Some drivers only viewed operational elements of the driving task as their responsibility:

"I felt no responsibility [for] driving just for emergency kind of like ready to press the brake in case somebody does something really irrational." [P5, OT]

Some drivers suggested that this assistance would extend beyond the takeover to directing decision for the driver in relation to the lane change manoeuvre:

"I think the assistance would, from the car would be telling you to pull off at the right time ... telling you where the space is [to pull in]." [P19, OT]

Most drivers assumed takeover requests were issues due to a critical incident or an automation failure, rather than approaching an operational boundary that can be planned for:

"I was trying to figure out why it said prepare to drive ... to see if there were any things that I needed to be wary of before I took control back... because there's a reason that it can't handle being in automated mode." [P10, OT]

This mental model of a transition being for an emergency seemed to influence what drivers expected to focus on/look out for during the transition of control. For example, looking for a specific hazard within the driving environment rather than general scanning patterns to raise their SA for a transition to manual driving. As a result, drivers expected the HMI to make them aware of a hazard, rather than thinking that they needed to build shared SA with the ADS by sampling the external environment over the course of the transition in preparation to resume the driving task.

"I had this in my mind that ... if there is any hazard or something the car will alert me, so I can use my phone and if something appears, comes up on the screen I'll just put my phone away and then get ready to take control." [P3, OT]

There was also a suggestion by one driver that the automation might request a takeover to test their level of attention, but again to a particular hazard within the driving environment:

"It might ask you to take over if you haven't taken control for very long, just to make sure that ... the driver is, is aware of his

surroundings ... if there is any hazard ... the car might be ... trying to test you.” [P

Other drivers highlighted that drivers needed to make checks and there was no need for the HMI to do this for the driver when they were taking over the driving task:

“If there was a vehicle close behind or you were approaching something then the car ... could tell you but I don’t really see the need because ... you’re taking over anyway and to me that’s the sort of things you should check anyway ... it might be beneficial to some people but ... I would check everything around me make sure I wasn’t going too fast and I was ready to keep the car at the same speed and direction as everyone else.” [P15, OT]

Another driver highlighted the potential issue of cognitive tunneling, suggesting that if the HMI highlighted specific hazards within the driving environment the driver may only focus on those areas and not carry out their own checks leading to them maintaining a low level of SA that could lead to errors.

“The problem with the vehicle highlighting dangers is that it then focuses all your attention on that danger and stops you from actually doing your own check ... it might be better to let the person take in all the dangers ... there’s a potential at least to force the person to focus on one thing to the detriment of their overall awareness.” [P20, BT]

Responses from the drivers highlights potential areas for concern relating to their expectations of who has responsibility for the driving monitoring task, in particular during the transition period between the TOR and manual driving mode engagement. Driver expectations regarding information provided by the HMI could influence the level of attention drivers pay to the driving environment prior to manual control. Low attention allocation risks a spike in workload when the driving task is resumed. In addition, focusing on one specific object or area in the driving environment or relying on the HMI to guide driver attention increases the potential for automation bias or cognitive tunnelling effects where the driver remains reliant on the ADS for their SA for the driving task. This highlights the need for careful HMI design to encourage appropriate scanning and sampling behaviours from drivers prior to taking back manual driving control. In addition, many of the comments highlight a need for appropriate mental model formation relating to driver responsibilities and a need for a joint cognitive systems approach that will encourage shared situation awareness of both human and machine agents. An in-vehicle HMI could provide a useful means to provide procedural guidance to drivers much like a pilot’s primary flight display (Trosterer, 2017). Behavioural or procedural instructions delivered via an in-vehicle HMI serves to provide drivers

with guidance both from a layered sense, in addition to other forms of training, but also in a tailored and pragmatic sense where guidance could be delivered in real-time as part of the vehicle ADS. It also provides opportunity for repetition and consistency to provide structure to novices and reminders to guard against complacency.

5.4.6.5.1 Use of Voice/Speech

A number of drivers in both groups mentioned the use of speech in relation to feedback or communication from the car. Drivers seemed to be expecting the car to talk to them in some way, even those drivers in the operational group who hadn't received the CHAT training:

"If it spoke to you somehow ..." [P8, BT]

"Potentially if there was a voice saying ... any other information ... that I need to be aware of or useful for me to engage differently." [P14, BT]

"Possibly verbal, um some kind of like ... the Knight Rider car, where... there's a voice actually telling you" [P5, OT]

Even when speech was not directly mentioned as the feedback modality, the feature description suggested a quality that would be associated with a verbal audio cue (as opposed to non-verbal alerts). Additionally, drivers tended to suggest that feedback should be designed to avoid causing a spike in workload or a startle effect:

"Some kind of audio cue as long as it was ... in a calm manner that ... wasn't information overload ...at that stage because ... someone may get flustered maybe not have ... loads of different cues all at the same time." [P13, OT]

A number of drivers highlighted a desire for some form of decision support feedback from the HMI system that emulated tools readily available in most modern vehicles, such as blind spot warnings and navigation systems were seen as particularly important as the drivers were unfamiliar with the road environment and didn't know where the motorway exit was. Spatial awareness and noting motorway signage seemed to something drivers were reliant on the vehicle to provide and were seen to influence driver decision making.

5.4.6.6 Trust Development and Feedback from the vehicle ADS

Driver trust appeared to be tightly linked to feedback from the ADS. Most drivers highlighted that they wanted more feedback from the vehicle, and how it would respond to elements in the driving environment, for confirmation that the ADS was engaged and working well.

"I might be curious to know more information ... feedback that ... I'm [the vehicle] doing my job with this level of accuracy, which might be reassuring." [P22, BT]

"It all looked like it worked as it should ... but ... the road was straight nothing really changed about the driving behaviour of the car so I wasn't sure ... I definitely was wary." [P23, BT]

"I was looking to see the car make some sort of action ... I was looking for some reassurance that it had control." [P4, OT]

Drivers were looking for evidence of how the ADS responded to other road users to help them form their mental model of how the vehicle's ADS operated and the capabilities and limitations of the system. This influenced how comfortable the drivers were in automated mode and whether they would need to respond to a critical situation. Drivers were told at the beginning of the study that the car's capabilities included OEDR, but drivers were looking for evidence during the drive that this was correct to provide reassurance and build trust.

"If a car pulled out in front of you, I presume it should be able to react to ... keep within the rules of the highway code and that sort of stuff." [P8, BT]

"I'm just seeing how the car reacts ... if any other drivers are ... changing lanes ... in the case that someone does cut me off, experiencing the dangerous side ... initially I was very wary because I didn't know how it would react." [P9, OT]

However, in time, even without direct evidence of the vehicle responding to other road users (due to limitations with the simulation software and what could be coded into the scenario), drivers reported developing trust with the vehicle's ADS and reducing their attention to the driving task.

"I trust it because it felt like a simple enough situation ..." [P22, BT]

"You think oh ok I don't really need to pay attention ..." [P12, BT]

Drivers also reported that a lack of feedback from the external environment was the reason their attention reduced. Their view was that their attention would have been drawn to changes in the driving environment and made them more alert.

"I think I got quite relaxed with it ... I noticed that no cars were coming on or off the carriageway which I think would make you more alert ... if ... there's cars coming in [on the junction] then you would expect the cars to your left to move ... it was ... starting to feel familiar." [P19, OT]

"I think there'd be a few cues as it got busy ... you'd start to see in your periphery ... more things moving at higher frequency and

you'd feel the car adjusting ... every time you felt one of those adjustments that would bring your attention back up and I think if that's happening all the time it will just keep ...my attention closer to the road." [P21, OT]

5.5 Discussion

The aim of the simulator study was to compare the CHAT behavioural training with a more traditional approach. It is possible to view those receiving the CHAT behavioural training as having been provided a more immersive and engaging experience compared to those receiving Operational training, and therefore a direct comparison is unfair. However, in practice, the training experience provided as part of the simulator study was designed to be comparable between both groups. It is also worth highlighting the following points to support the selected approach:

- **Authenticity.** The Operational training is a true reflection of the level of detail and approach offered by current user manuals for automated systems. In practice, it is expected that a lot of new users will not read their user manual in great detail, if at all. However, during the study, dedicated time and specific instruction was provided to drivers, to read through the manual in detail and this was followed by a test drive scenario to re-enforce the learning.
- **Content.** The CHAT behavioural training did not include any functional details about how to use the system, typically found in user manuals. This cohort *only* learnt how to use the automated system via the test drive scenario, and did not have sight of any details regarding the functionality of individual features such as autosteer/ACC. In other words, they were only provided with information on the overarching limitations of automation at L3 and what this meant for their interaction with the vehicle, but no specific details on how the system operated. The CHAT behavioural training also focused on the wider context of automation to aid understanding, i.e. 'automation policy' and 'automation principles', which was absent within the functional approach of a user manual.
- **Timing and Trainer Involvement.** Both approaches lasted the same amount of time and neither involved face-to-face interaction with a trainer, albeit the CHAT behavioural training made use of expert commentary to aid delivery. Even so, this still relied on the motivation and understanding of the learner to self-administer the training (i.e. they were required to select and playback episodes and move through the presentation at their own pace).

- **Training Assessment.** All participants were asked to evaluate their training experience using a questionnaire adapted from a classroom engagement study by Wang et al., (2014). Results from the training evaluation questionnaire show that there were no significant differences between groups in terms of engagement, usefulness and attitude. This indicates that both approaches were comparable from the perspective of the participants' experiences.

5.5.1 Subjective ratings

Whereas trust ratings were comparable between groups at the start of the study, the subjective data shows that participants who received Operational training expressed higher trust-in-automation and higher intention to use SAE-level 3 automation after their experience and felt their situation awareness (SA) was as good as those receiving Behavioural training. In contrast, those in the Behavioural group felt greater time pressure (temporal workload) associated with the demands of the task (maintaining awareness, sharing control, keeping on or in-the-loop etc.), and indicated no differences in their trust and intention to use the automation after the drive.

On face value, the outcomes associated with the Operational group may seem to be more favourable; after all, it is important that people trust future technologies and intend to use them. Moreover, lowering workload may, on initial inspection, appear desirable. However, the concern is that at intermediate levels of automation, the driving task is shared, and the driver therefore retains responsibilities and has specific tasks to undertake, even when the vehicle is notionally in control. Consequently, trust must be calibrated appropriately to avoid the driver over-trusting or over-relying on the technology. Additionally, achieving a level of workload commensurate with the task at hand (i.e. not removing it completely), is not only favourable, but also a prerequisite to help drivers remain alert and maintain awareness (Young and Stanton, 2007). It is also expected to support the accurate, ongoing calibration of trust. On this basis, especially when considered alongside metrics, such as the number of drivers who were aware of the hazard vehicle, it is posited that drivers in the Operational group may be forming opinions and making decisions based on limited knowledge or poorly constructed mental models, leading to overly optimistic assumptions about the capability of the automation, and unsubstantiated judgements about the level of trust they should place in the system. Such factors have already been highlighted as possible deleterious consequences of vehicle automation (Seppelt and Victor, 2016; Kyriakidis et al., 2019). These are exactly the elements that the Behavioural training aimed to address. Indeed, trust ratings made by drivers in the Behavioural group were unchanged after the experience.

5.5.2 Visual behaviour

During the study, drivers provided subjective ratings of their SA using the situation awareness rating technique (SART) (Taylor, 2017). It is noted that it can be difficult to guarantee that the SART scale items have been interpreted and applied correctly, particularly in the context of simulated driving. There is also an inevitable delay before responses are collected, which can impact on the ability of participants to accurately recall their experiences (Salmon et al., 2009). In contrast, 'freeze-probe' techniques, such as the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988), provide the opportunity to question participants at the point of criticality. However, these may still be unreliable and subject to bias in certain contexts. For example, obscuring the road scene and asking participants to recall precise details of the type and location of all other road users, etc., fails to acknowledge the dynamic nature of the driving task, and the relative levels of risk presented by different road users. For instance, a driver might be expected to direct greater attention towards a vehicle displaying unusual behaviour (e.g. unexpected braking, swerving or tailgating their own vehicle) and subsequently reduce the relative level of attention they direct towards other apparently 'safer' drivers. As such, when asked, they may be able to provide a precise estimation of the location of the former, but not necessarily the latter road users. To overcome some of these limitations, it has been suggested that a range of measures should be employed to assess SA from a number of perspectives (Salmon et al., 2009). In particular, those that show the driver is making an active attempt to engage with the driving environment and key elements within it. In this context, eye movements, such as glances to mirrors, have been shown to be one of the most perspicuous activities (De Winter et al., 2019). Mirror checks are considered essential for safe driving and help drivers to maintain (or rebuild) SA (Li and Busso, 2013). This is particularly the case in the context of automated driving, whereby drivers may be engaged in NDRTs: checking their mirror, or the roadway, demonstrates a clear and deliberate attempt to disengage from these activities (if only temporarily) and rebuild their awareness of, or to remain engaged with, key elements in their surroundings during automated driving.

The results demonstrate that drivers in the Behavioural group made significantly more mirror checks during automated driving and when asked to prepare-to-drive than those who received Operational training. In practice, drivers receiving CHAT behavioural training carried out over 30 additional mirror checks compared to those receiving Operational training (on average, 47 per person compared to 16.8, respectively), even during this relatively short journey. They were also more likely to see the tailgating hazard car, which appeared at the same time as the

takeover request was issued – ten out of the eleven drivers (over 90%) in the Behavioural group saw the hazard car, compared to only three of the thirteen drivers (23%) in the Operational group. Given that four drivers in the Operational group made at least one glance to their rear-view mirror during the transition of control, it can be assumed that one of these drivers looked in their mirror but failed to notice the hazard car. Inattention blindness, or ‘look-but-failed-to-see’ errors have been highlighted as a major cause of road traffic incidents, particularly involving more vulnerable road users, such as cyclists and motorcyclists (Herslund and Jørgensen, 2003). This detail was confirmed in the post-study interview and provides further support for the provision of a specific operating procedure, such as CHAT, to encourage conscious processing of the checks and assessments drivers must do at this point in the drive.

One of the notable problems highlighted in the observation study described in chapter three was that drivers remained actively engaged with their NDRTs after receiving the TOR instead of preparing to drive. Moreover, the amount of time they continued to remain engaged with NDRTs increased over the week, with drivers choosing to continue with these activities for longer each day, even during the prepare-to-drive period. It was also noted that these drivers tended to make an initial, perfunctory glance to the road when asked to prepare to drive, but then immediately returned their attention to their NDRT, choosing to prioritise this instead. In the current study, this period of ‘shared attention’ was defined as the time from the driver’s first attempt to re-engage with the driving scene, indicated by their first driving-related glance (e.g. to a mirror or the forward road scene) immediately after the TOR was issued, until their final NDRT-related glance or interaction. It is interesting to note that drivers who received CHAT behavioural training demonstrated a shorter period of shared attention, putting aside their NDRTs earlier than those who received Operational training. On average, drivers in the CHAT group started to re-familiarise themselves with the driving scene (i.e. start of shared attention) 14 seconds earlier and had completely discharged their non-driving related activities 23.5 seconds sooner than Operational drivers. This naturally provides more time to enable drivers to re-familiarise themselves with driving.

Mirror glances made by drivers following the takeover of manual control were categorised as serving one of two purposes. Initially, glances were used to build general SA (allowing drivers to contemplate changing lanes), and subsequently they were used to maintain awareness during the lane change manoeuvre itself. Drivers were observed typically making several mirror glances after resuming manual control, which were used to build awareness of their surroundings, but did not precipitate the action of changing lanes. The point at which it was deemed the driver had confirmed that it was safe to change lanes

(based on observations during video analysis, e.g. hands moving to the indicator; and driving performance data), mirror checks were thereafter considered employed to maintain awareness during the manoeuvre itself. It is interesting to note that drivers receiving the CHAT behavioural training not only made significantly more mirror checks after resuming manual driving and *before* starting the lane change, they also made significantly more mirror checks during the lane-change manoeuvre itself, even though, arguably, they were already in possession of greater awareness of their surroundings (given the higher number of mirror checks during automated driving) than drivers in the Operational group. It could be interpreted that this behaviour may reflect (and also require) additional workload and effort. However, it is also indicative of a more cautious approach, which is arguably more appropriate given that drivers may have been 'out of the loop' with respect to the driving task for an extended period of time.

5.5.3 Out-of-the-Loop, On-the-Loop and In-the-Loop

Attendance to an NDRT (including thought, i.e. 'mind not on driving') during automated driving induces an OOTL driver state (Merat et al., 2019). It could be argued that without appropriate training or HMI support to improve SA in a timely way, there is a residual impact on driver SA and vigilance well after explicit shared control has visibly ceased. For example, the lack of mirror glances conducted by drivers from the Operational group during the transition period demonstrated a failure to observe safety critical areas of the roadway. This, in turn, negatively impacted their visual search and selective attention performance in perceiving the tail-gating car and putting them at increased risk of inattentional blindness (Mack and Rock, 1998), as evidenced by the driver who 'looked but failed to see' the hazard car even though they had actually looked in the mirror. In other words, during the transition phase from automated to manual driving, drivers in the Operational training group could be said to be 'on the loop' in relation to the innermost loop in Merat et al.'s (2019) multi-level control in driving model, but remained 'out of the loop' at both the middle and outermost loops (see Multi-level control in driving, including situation monitoring at each level). In contrast, those drivers receiving Behavioural training could be said to be 'on the loop' (for all loops) during the transition, actively monitoring the driving situation and attending to multiple on-road regions, which resulted in a greater number of mirror glances in relation to the basic vehicle motion control and during the planning and execution of the lane change manoeuvre.

5.5.4 Driving performance

Following the handover of control, participants had been instructed to move into lane 1 and exit the dual carriageway. Given the presence of other road users in lane 1, this meant that they were required to find a

suitable space between vehicles already travelling in lane 1 and move safely into this before leaving the road. However, the other road users in lane 1 were not necessarily in the same positions relative to the ego vehicle for each participant, especially as drivers were able to choose to curtail the 60 seconds prepare-to-drive notification and resume manual control sooner, if so desired. Consequently, there was no 'perfect', predefined manoeuvre with which to compare driver groups. It is therefore difficult to draw robust conclusions from the driving performance data. Moreover, in the chosen scenario, which was specifically selected to ensure drivers were required to make the type of manoeuvre that might be expected following the transition of control, behaviour may be interpreted in different ways. For example, a swift move into lane may be due to the fact that the driver already had good awareness of their surroundings and were confident that it was safe to change lanes immediately; equally, it could be because the driver *thought* it was safe to change lanes, but in fact it is not, i.e. it is precisely their lack of awareness that has precipitated the manoeuvre.

Consequently, it is not possible to draw absolute comparisons between groups in the current study. As such, the lack of significant differences in driving performance measures is not of particular concern. Rather, it is suggested that this is simply symptomatic of the nature of the driving task immediately following the handover of control, and an artefact of the naturalistic design of the scenario. Moreover, no specific technical takeover skills were imparted during the training, and drivers were only in attendance on a single occasion. Nevertheless, it is suspected that different results and driving performance may be observed in a different post-takeover scenario, which will be taken forward and explored in follow-on studies for the thesis. However, it is noted that some caution will need to be applied in the selection and analysis of a post-automation scenario. For example, other related work that has employed a hazard situation immediately after the transition of control, such as a braking vehicle ahead, has found that drivers' response to a potential collision event such as this is often reflexive, typified by sudden, emergency braking (Zeeb et al., 2016; Gold et al., 2018). Therefore, drivers may be successful in avoiding the hazard incident presented to them. However, it does not necessarily indicate that they are fully prepared to control their vehicle and are able to maintain the relevant capability to manage the driver task on an ongoing basis.

5.5.5 Benefits of the CHAT Standard Operating Procedure

The thematic analysis of the semi-structured interviews highlights further evidence in line with the literature (McDonald et al., 2018; Casner and Hutchins, 2019), that many drivers do not accurately understand the capabilities and characteristics of level 3 automation.

There was also a demonstrative lack of understanding from the operational training group regarding the required level of control and engagement with the driving task during a planned transition from automated to manual driving modes. The emerging themes provided further support to the necessity of training and guidance interventions for drivers of these future vehicles. Also, adding weight to the challenge regarding the assumption that current methods of training for experienced drivers, such as the use of vehicle operating manuals, are a sufficient way to guide required driver interaction.

The CHAT standard operating procedure used as a core part of the Behavioural training was developed using a combination of design principles informed by the implications of long-term memory on learning. Additionally, a central part of the design of this training tool was the importance of meaning and structure in the efficient application of skills-based knowledge to the nuanced and dynamic driving context. Drawing upon the learning strategy of 'proactive observation' (Castro et al., 2016), the aim was to highlight the importance of improving SA in preparing to drive following a takeover request (TOR). It was designed to draw the trainee's attention to the number of events and (typically automatic) actions relevant to making proactive and reasoned decisions that are required **before** taking over physical control of the vehicle to enable safe and effective driving. The CHAT procedure gave trainees a template for how to actively scan a takeover scenario for relevant objects or events, and expert commentary to guide their attention and provide immediate feedback to the task. In this way, the learning approach and procedural design aimed to motivate the driver in the uptake of desired behaviours, and that this would facilitate them in being 'on-the-loop' during transitions of control.

Comments from drivers in both conditions supported the view that, without adequate training or guidance, drivers find it hard to remember, or even recognise, the importance of attention allocation to the external driving environment during planned transitions of control (Endsley, 2017a; Carsten and Martens, 2019; Mole et al., 2019). As was inferred in the findings from the observational study in chapter three and well as behaviour in the operational group in this study, drivers may glance at the forward road scene following a TOR, but these glances are more likely to be fleeting, narrow in focus (i.e. limited to the forward rather than surrounding road), and lacking the motivation and purposeful goal of raising and recalibrating their SA prior to taking control of the driving task. The CHAT acronym was designed to attend to several principles relating to supporting learning, memory and understanding (Wickens et al., 2014), including the semantic association with our understanding of the importance of communication in collaborative partnerships. The selection of the word 'chat' was intended to inspire and reflect the process of conversation that facilitates mutual understanding, and

efficient and effective collaboration and decision making when working towards a shared goal. Both the objective behaviour results and interview comments from the CHAT training group point to an immediate potential benefit of enhanced knowledge and subsequent influence on driver behaviour during the transition period. Driver feedback and the results from driver behaviour suggests that the CHAT training was effective in raising the need for glance behaviour that provided ongoing sampling of the surrounding external driving environment (e.g. including mirror glances), to support SA recalibration and decision making into the conscious thought processes of drivers.

Drivers in both conditions acknowledged the need to take on responsibility for the monitoring task during the transition period, effectively becoming 'on the loop' of control. However, this was accompanied by the recognition and concern relating to retention of training and automatic processing of skill-based behaviours typically involved in the driving task, specifically pointing to the concern that a one-off training module maybe inadequate to provide guidance for required behaviour on an ongoing basis and in real-time when drivers are performing a transition of control. This was in addition to the challenge that a formal driver training approach may not be appropriate for all types of drivers, for example, experienced as opposed to novice drivers. This highlights the need for a tailored and layered approach, for the purposes of flexibility and redundancy, to ensure appropriate interaction is developed with drivers of these future vehicles. This points towards the use of operational procedures and HMI design as potential methods for delivering driver guidance and support in an ongoing, real-time and accessible way.

5.5.5.1 Application in Human-Machine interface (HMI) design

In joint cognitive systems (i.e. where the human and 'intelligent' machine each play a role in a shared task), such as driving automated vehicles, both the human and the system need to collaborate to deliver safe and comfortable driving. The main means of communication between the vehicle and human to facilitate this partnership is the human-machine interface (HMI). A key role of the HMI is to help the human driver understand what is expected of them in terms of monitoring and active interventions (Carsten and Martens, 2019). As system design affordances (i.e. the qualities or properties that define how something should be used) (Norman, 2013) can lead to misuse or incorrect interactions with vehicle automation (Merat et al., 2019), so the HMI plays an important role in supporting appropriate driver behaviour. Following ecological design principles (Rasmussen and Vicente, 1989), the goal of an HMI should be to support the efficient, rapid and accurate recall of desired skill-based behaviours and avoid

the need for drivers to employ reflective interpretation of information that could slow down required responses in time-critical scenarios. The CHAT procedure was designed with these key points in mind. It presents a potential design template for use in HMI design that has the potential to provide an effective countermeasure to human performance problems relating to OOTL driver state induced by automated driving at level 3 automation and exists in a format preferable to decision tools or remedial skills training (Carsten and Martens, 2019).

Additionally, driver interview comments relating to an anthropomorphic, 'voice' element to HMI design suggests a desire for a conversational approach to communication and feedback from vehicle systems and displays to assist with the transfer of the driving task from automated to human control. Evidence for the potential in this type of approach has been well established within the driving context (e.g. Large et al., 2017b). There is an argument for providing drivers with hazard detection display technology, for example in the form of augmented reality, that can draw driver attention to specific objects within the driving environment. However, there is research evidence to suggest that this approach may not provide suitable guidance for appropriate glance behaviours and sampling patterns, as it can focus drivers' attention on one aspect of the driving environment to the detriment of others (e.g. Eriksson et al, 2018). Recent research by Clark et al., 2024, that analysed the speech interactions between two human drivers simulating a planned control transition, provides useful insight into the potential for conversational HMI designs to encourage more joint operations between human and vehicle.

5.5.5.2 Standardisation and Versatility

In addition to its proven effectiveness at motivating behavioural change in situational monitoring, the concept of CHAT has further applied benefits in relation to standardisation and versatility that make it an attractive concept for further development. First, the CHAT concept provides an opportunity for a standardised approach to an, arguably, essential operating procedure relating to transitions of control in level 3 automated vehicles and is independent of differences in system design. Its distinctive and memorable design and succinct format make it practical for use either within HMI solutions, or, if vehicle manufacturers are averse to using the concept within their designs for aesthetic reasons, it can be applied within public safety or marketing campaigns (similar to the 'THINK!' road safety campaign (Department for Transport (DfT, 2020)). Repetitive use, facilitated by either of these approaches, should aid learning. In this case, the CHAT procedure moved from declarative to procedural knowledge and encouraged transfer of training (Krampell et al., 2020). The use of CHAT within HMI design offers a way to improve SA and attention in a timely way as communication via the HMI can promote efficient scheduling of situation monitoring tasks

with the driver, a key goal in procedural design (Degani and Wiener, 1997). Additionally, the explanatory training related to CHAT and the guiding philosophies of automation principles, policies and procedures is versatile enough to be either integrated into training for new drivers or delivered on a stand-alone basis to experienced drivers at point of sale or hire of a level 3 automated vehicle.

5.6 Limitations

Overall, results suggest that behavioural 'CHAT' training had a positive influence on tactical level task performance during automated driving and following a transition to manual driving. It is also possible to infer that the early engagement in re-building SA demonstrated by drivers in the Behavioural group led to more informed and measured decision making in relation to the lane change manoeuvre. However, caution must be used when interpreting the success of these results. First, the effect on knowledge retention and maintenance of desired behaviour has not been tested within the present study. Second, success of any proposed driver training intervention will depend on the willingness and compliance of drivers to complete it and the relevant bodies to facilitate, finance and regulate its development and management. These challenges make a valid case to investigate ways of integrating the key concepts used in this Behavioural training intervention into a technological design solution. In particular, as the training intervention did not aim to introduce any new skills but rather apply those that experienced drivers will already have, it is envisaged that these would represent a specific operating procedure to support the new automated driving context. Thirdly, it is worth noting that while it is possible to infer certain conclusions through the cross referencing of objective data, such as mirror glances, with events in the scenario (e.g. movements of a hazard car or lane changes) and verbal comments through interviews, individual thoughts are not directly observable and themes extracted from qualitative data are subjective opinions. For example, we cannot objectively say that if a driver looked in the mirror during the time a vehicle was behind them that the driver perceived that the car was there. Equally, it is not possible to remove all bias from a researcher when seeking themes from qualitative data. However, this does not make the conclusions drawn from this study invalid, indeed the findings can be used to inform future interventions that can be further tested in future studies to increase the validation of these foundational results.

It is also worth recognising that the results as presented relate to the specific set of controlled conditions to which participants were exposed, the most significant factor being that the study took place within a medium fidelity driving simulator. To address such concerns, the study aimed to ensure that the experience was as realistic as possible. A naturalistic approach, as far as was practicable, was taken to reflect

future expectations related to the introduction of L3 ADS. For example, participants could select their own NDRTs and choose when to disengage automated driving following the TOR. Nevertheless, there remain unavoidable, inherent limitations associated with research conducted in a driving simulator, relating to factors such as the fidelity of the experience, the absence of risk from harm if a collision occurs, and so on (Caird and Horrey, 2011). For these reasons, caution should be applied when drawing absolute conclusions and relating these to on-the-road behaviour. However, it is proffered that the *relative* validity of the study findings can be reported with confidence, alongside the *absolute* value and potential benefits of behavioural change motivators, such as CHAT. Fundamentally, in the exploration of future technology that is not yet available for public consumption, particularly safety critical systems that present a potential risk to participating drivers, it remains an inevitable necessity to use simulated experiences and approaches to investigate their impact on human performance and safety. Finally, it is worth highlighting the imbalance in gender for the participants of this study. Strategies will be used for future simulator studies to increase the number of female participants taking part.

5.7 Conclusions and Future Directions

This chapter described a driving simulator-based study that was designed to evaluate the effectiveness of a proof-of-concept, behavioural training intervention. The training took the form of a specific standard operating procedure (SOP) based on the principles CHAT (Check, Assess, Takeover) that was designed to be applied during the transition of control, which was tested during this between-subjects driving simulator study to demonstrate its immediate and tangible benefits. Interestingly, drivers who had not received the Behavioural training ('the Operational group') judged aspects of their own behaviour and performance favourably, as evidenced by high ratings of trust; the belief that they were equally well-informed about the driving scenario (i.e. the same level of SA as the behavioural group); low ratings of workload; and the fact that they remained engaged with their NDRTs for much longer (suggesting a high degree of confidence in their ability to takeover at short notice).

Based on the findings from the current study, it is argued that the mental model of these drivers was likely to have been incorrect from the start. Therefore, the approach they took and the judgements they made were applied on an inconsistent and ill-informed basis. This may be adequate in some situations, and their performance during the transition of control may even be deemed "good enough" during the study, though clearly not optimal. However, on the road, where the dynamic environment is typically unpredictable, this may not be the case. The concern is that because of their lack of knowledge and awareness,

these drivers may make overly optimistic assumptions about the capability of the automation (and of themselves), leading to unacceptable or dangerous behaviour. Moreover, for these drivers, their mental model will arguably only be updated or corrected by chance; for example, if they were directly involved in a hazardous situation.

The CHAT behavioural training aims to ensure that drivers are aware of what they should or could be doing to help prepare themselves to take over control following a period of automated driving. It also encourages drivers to remain engaged with the tactical and strategic elements of the driving task during periods of automation and the transition of control and avoid focussing solely on the operational or control tasks during the takeover. It is therefore notable that ratings for trust and intention to use did not increase after the experience amongst those who received behavioural training.

The improvements in visual behaviour are abundantly clear. Drivers in the Behavioural group were more likely to conduct checks to the internal and external driving environment during the transition from automation to manual mode than drivers who received the Operational training. They were also more likely to demonstrate positive visual behaviours relating to tactical and strategic level tasks during the automated driving mode than their counterparts. They were arguably therefore better prepared to successfully perform a manoeuvre shortly after transition from automation to manual control than drivers who had received the Operational training.

Nevertheless, it is recognised that the effect on knowledge retention and maintenance of desired behaviour is yet to be tested. Moreover, as already highlighted, the success of any proposed driver training intervention will depend on a number of interrelated factors. These factors are outside the scope of the current study. The next phase in this research will focus on using the novel CHAT procedure to inform the design and development of a human machine interface to support the uptake and, importantly, the maintenance of the desired driver behaviours.

The final section of work for this thesis (chapters six to seven) intends to extend these research findings by seeking to further validate the overall necessity for behavioural intervention in non-critical, planned control transitions and explore ways to improve the evaluation of the effectiveness of the CHAT SOP on supporting driver transitions between control loops. There is a need to explore whether formal training is the best approach for experienced drivers, or if the introduction of the CHAT principles into manufacturer guidance or in-vehicle technological solutions (e.g. HMI design) is sufficient to mitigate against the performance issues identified by previous research. CHAT may, therefore, present an opportunity to improve driver safety at

level 3 automation without the need for formal training interventions or graduated licences, reducing disruption to the roll out of such vehicles without compromising on safety.

6. Design of the CHAT Human Machine Interface

6.1 Introduction

This chapter outlines the design of the CHAT HMI and talks through the theory that influenced its development. In particular, it is highlighted how research into the use of naturally occurring human-to-human speech data for the design of conversational user interfaces (CUIs) shaped the way audio voice prompts were structured to support the mental model of a collaborative team.

6.2 Background

The physical and cognitive disengagement of drivers with the driving task during level three automated driving negatively impacts their understanding of the task demands and retention of critical information for effective driver capability (Gugerty, 1997; Endsley, 2018). Therefore, if drivers are to safely transition to and resume effective execution of the driving task, they must have sufficient time and resources to recalibrate the level of SA and workload required to handle the dynamic nature of the driving environment (Carstens and Marten, 2019; Gold et al., 2013; Endsley, 2017a).

A driver's mental model of the current driving situation dynamically controls how they allocate their attention, guiding (or not!) their search for new information (Endsley, 2017a). In-vehicle HMIs offer a potential solution to influence drivers' allocation of attention to competing information (e.g. between NDRTs and the primary driving task) during automated driving (Endsley 2017). Effective HMI design could offer a significant support to raising and recalibrating driver SA and workload as well as forming and reinforcing driver mental models relating to planned control transitions (Carsten and Martens, 2019; Endsley, 2015). Key information for the development of driver SA must be identified and included in the system design. However, this information must be presented at a manageable level of complexity, which facilitates the development of drivers' accurate mental models and the effective use of driver's own senses (as opposed to information cueing), to align control actions and maintain safe system operation.

It is, arguably, unrealistic to expect drivers to monitor the driving environment during ADS engagement, especially as this would defeat many of the proposed advantages of using such systems. Therefore, to effectively manage planned control transitions, drivers need a level of information and guiding principles to help them understand and undertake actions to satisfy transition readiness requirements ahead of resuming manual control. Checklists are habitually used within the aviation domain as an effective strategy to provide routines to guide

rule-based behaviour for required competence in specific procedures (Trosterer et al. 2017). As such, they provide an interesting avenue to explore in the context of PCTs. In particular, if the ODD is known then it should be possible to apply a shared driving concept to plan and delegate different driver involvement, in accordance with predictable and transparent variations in system capability (Trosterer et al., 2017). For example, on long motorway journeys, it is reasonable to anticipate the need for driver involvement when exiting the motorway. This would enable advanced system communication of a required transition giving drivers notice to plan and adjust their SA, allowing for NDRT engagement during periods of automation and recalibration of SA ready for manual resumption of control. Taking inspiration from interviews with pilots, Trosterer (2017) suggests that a primary driving display, much like a primary flight display, could be used as an efficient way to communicate information, such as a checklist of actions required of the driver prior to taking over control of system operations. Such a strategy offers a potential way to support recalibration of driver's cognitive engagement with the driving task. By breaking down the tasks required during a PCT, it could encourage effective driver behaviours, to mitigate potential errors borne from a mismatch between task demands and driver capability following the resumption of manual driving.

This research proposes that the CHAT standard operating procedure (SOP) offers a checklist template for use in HMI design that has the potential to effectively support drivers during PCTs. Furthermore, it is suggested that the delivery of the CHAT SOP in an HMI format is preferable to decision tools or remedial skills training, as it provides consistent, repeatable and real-time support for efficient, rapid and accurate recall of required behaviours (Carsten and Martens, 2019). A combination of emergent themes drawn from driver interviews from the training simulator study; a review of the literature on HMI guidance for human-automation interaction; and a recent proposal on the use of speech acts in virtual assistants (Clark et al, 2024) were used to develop a 'CHAT HMI'. As part of this work, consideration was also given to the development of the final simulator study design and effective evaluation of key concepts such as driver situational awareness.

6.3 HMI Design for Control Transitions

The physical and mental decoupling from vehicle operations at L3 automation poses a significant challenge to driver attentional allocation and subsequent SA (Carsten and Martens, 2019; Endsley, 2017a; Mole et al., 2019). However, according to Carsten and Martens (2019) although this disengagement inevitably leads to a reduction in driver SA and attention during periods of driving automation, it is not necessarily unsafe. Rather, if the ADS reliability is appropriately calibrated with

control requirements and can be communicated in advance, then use of a well-designed HMI system to support the timely recalibration of driver SA and attention will provide adequate system design (Carsten and Martens, 2019). However, the weight of the challenge should not be underestimated, especially as, due its dynamic nature, drivers will potentially be faced with a dramatically different driving environment at the point of a required transition (Mole et al., 2019). Additionally, as outlined earlier, the unpredictable temperament of traffic conditions means that drivers may find themselves confronted with a sudden, urgent, or even critical situation that exponentially increases the task demand during or in the moments after manual driving has been resumed (Teh et al., 2014). Both contexts have been shown to increase driver workload and the probability that error will lead to an accident (De Waard et al, 2008; Brookhuis et al, 1991). The contextualisation of potential task demands, both during and following a PCT adds further emphasis to the importance of appropriate calibration of SA and attention to support safety and performance. Despite this, there remains a tacit assumption that as the level of control that can be devolved to the ADS increases, the task complexity, and therefore training or guidance required for the human driver decreases (Cummings et al., 2019).

As system capability increases to SAE L3, so does system reliability and, subsequently, the ability to increase the level of prediction for driver fallback requirements (Trosterer et al., 2017). In other words, the capability of the ADS to control the task of object and event detection and response (OEDR) should mean a shift in focus away from the current predominant paradigm of failure-induced control transitions towards planned or routine transfers of control. This would involve an associated movement of driver responsibility from a critical response requirement towards a more managed and timely response to a planned control transition due to the boundaries of the ODD being reached (Carsten and Martens, 2019). As already highlighted, although this increase in system reliability is positive for driver comfort and trust, the impact on driver performance during periods ADS engagement makes strategies to countermeasure issues of operator SA, attention and workload even more important. However, given the increased predictability of control transition requirements, this presents an opportunity for the use of HMI design as a suitable intervention to guide desired driver behaviour prior to resumption of manual driving control (Trosterer et al., 2017).

To date, PCTs have not been the predominant focus of the research into effective control transitions (Eriksson and Stanton, 2017b). However, factors such as system transparency, the coordination of driving task activities, and shared control between the automated system and the human driver have been highlighted as key elements

that need to be investigated and could be supported through appropriate HMI design (Carsten and Martens, 2019; Eriksson and Stanton, 2017a; Endsley, 2018). The concept of 'shared control' is often used to reflect the collaborative and coordinated requirements of the driving task in automated vehicles (Flemisch et al., 2012; Large et al., 2017a; Banks et al., 2019). It draws attention to the critical interdependent nature of the roles the automation and human agents play within a joint cognitive system; in particular, in relation to the challenges of shared situational awareness when operating a safety critical system within a dynamic environment.

Eriksson and Stanton (2017b) purport that a key purpose during the control transition phase should be to ensure the agent receiving control (in this case the human driver) has a current model of the system state, is aware of any significant changes to the state, and has the necessary knowledge to perform duties, including the ability to anticipate future events. Co-ordination of this kind of joint activity (as well as the driving task as a whole in the context of automated vehicles) depends on the ability to predict, with reasonable accuracy, the actions of the other agents involved in the team (i.e. the automation and the human driver) (Klein et al., 2004a; Klein et al., 2004b). These actions include an accurate understanding of features such as the level of difficulty of required actions and the time and skill required to complete actions (Klein et al., 2004b). A critical element to achieving this 'inter-predictability' is to establish mutual beliefs, knowledge and suppositions between the system agents, or 'common ground' (Klein et al., 2004b; Eriksson and Stanton, 2017b;). Establishing common ground, much like the formation of an accurate mental model, will ensure system transparency. An HMI system designed to provide co-ordination and common ground to the human-automation system, alongside timely and task-relevant feedback, would reduce the cognitive burden on the receiving agent (i.e. the driver) to determine the system state and whether their behaviour matches the expectations of the joint cognitive system (Norman, 2013; Eriksson and Stanton, 2017b).

Johnson et al. (2014) proffer that joint activity requires observability, predictability and directability (OPD) between team members to support interdependence. Carsten and Martens (2019) use these as a basis for their HMI principles for system transparency in automated vehicles. They highlight that, as the main means of communication between the automation and the human driver, the HMI plays a key role in facilitating the partnership and joint activity required for the coordination and shared control of the driving task. Careful HMI design is required to support the development of driver mental models that set appropriate expectations of driver responsibility and supports desired driver behaviour. Carsten and Martens (2019) emphasise the importance of predictability in automated system design to minimise unexpected

events that impact driver requirements, i.e. automation surprises. In particular, the critical requirement to ensure that system transparency and timely information is provided to the driver regarding the need for planned driver intervention. Using Johnson's (2014) OPD requirements, they outline that the HMI should allow the driver to understand or detect system status (observability), plan their own actions according to the observable, understandable and reliable actions of the system, and influence or be influenced by both the explicit and implicit commands provided by the HMI to the driver (or vice versa) to come to the best joint performance. They also add the additional requirement of timeliness, which relates to the provision of information or warning to the driver early enough that they can take proper action. It is noted that in the context of L3 automation this requirement will need to take account of the driver's out-of-the-loop driver state.

A final core HMI principle highlighted by Carsten and Martens (2019) work relates to the provision of psychological comfort to the human driver through the reduction of uncertainty and stress. They highlight the comfort that may be offered to the human users of automated systems by allowing them the possibility of doing something else and using their time in a valuable way. System reliability at L3 is designed to offer this. However, this will only be a reality if the system offers appropriate levels of trust, transparency and predictability that provides them with a balanced state of wellbeing. Carsten and Martens (2019) suggest that the way the vehicle behaves and the information it provides, should communicate to the human operator that the automation is in control, and the vehicle's behaviour is in accordance with predicative rules and explainable behaviour.

HMI interventions within automated vehicle research have tended to focus on warning systems that alert drivers to a TOR and direct them to take physical actions, such as placing hands on the wheel and eyes on the road. The overarching performance goal of these devices is to produce the fastest takeover response time from drivers (see Eriksson and Stanton, 2017a for a review). However, this approach is arguably more suited to the paradigm of emergency control takeovers, which depend upon fast and effective response from the driver in a time critical context. In this thesis, it is argued that a more measured approach could (and should) be taken with planned transitions that focuses on ensuring drivers are given adequate time and resources to safely take over control of the driving task.

More recently, Ericksson et al. (2018), conducted a research study to assess HMI systems that provide information support to drivers during the transition of control, to enhance the predictability of the situation. Each HMI design evaluated offered different levels of support, in accordance with the stages of Parasuraman's (2000) levels of automation framework. Four HMI designs were tested, these included a

baseline HMI that provided a vibrotactile warning of the TOR with no visual support, and three HMI designs that provided drivers with different types of visual feedback at the same time as the TOR. The visual feedback consisted of an augmented reality overlay to highlight either hazards in the environment (information acquisition); information on the availability of a gap in the adjacent lane that the ego vehicle could pull into (information analysis); or arrows indicating either an action for the driver to change lanes or brake (decision selection). Results showed that the HMI design did not affect driver reaction times to the TOR, but the visual support provided by the decision selection HMI did provide improvements to the success rates of manoeuvre selection (either to change lanes or brake). However, a key point of note was in relation to the use of information acquisition support in the form of information cueing. Not only did this HMI design fail to improve driver decision making, it also increased unnecessary braking and yielded the lowest success rate of all conditions (including the baseline) in the lane change scenario.

These conclusions highlight the potential detrimental effects of support systems, which have already been observed within the aviation domain, such as skill degradation (Haslbeck and Hoermann, 2016), complacency (Parasuraman et al., 1993), or automation bias, in the form of errors of omission or commission. These forms of error are respectively caused by a failure to act due to lack of system information (Stanton et al, 1997), or implementing an incorrect action based on system information coupled with a failure to check direct information (Skitka et al., 1999). The performance improvements found by Eriksson et al., (2018) add support to the views of Endsley (2015), who specifically advocates for the use of carefully designed, automation interfaces as a tool to directly improve SA by influencing drivers' ongoing allocation of attention, rather than the use of decision tools or remedial skills training. However, their full results add a cautionary consideration in relation to HMI design approaches. This is reflected in Endsley's (2017) view that HMI systems should avoid the use of information cueing and instead should use approaches that encourage operators to use their own senses more effectively. It is worth noting that the persistence of the issue of automation bias has been highlighted across all drivers regardless of experience (Mosier and Skitka, 1996).

A recent study by Clark et al., (2024) highlighted how research in the field of onboard virtual assistants, featuring conversation user interfaces, offers a novel and flexible approach to HMI design. The use of virtual assistants as a form of informed, attentive team member that can provide immediate and information-rich communication in the context of driving are well established (e.g. Large et al., 2017b). This design approach offers a potential way to explore how situation

information should be delivered between agents to best coordinate the recalibration of SA and attention and support desired driver behaviour during planned transitions of control.

In their research, Clark et al. (2024) employed a novel experimental design, which aimed to use human-to-human communication to explore how natural language and coordination is used during the takeover task. Pairs of participants took it in turns to make a human-automation driving 'team', where one participant took on the role of the automated system and the other, the role of the driver. In their respective roles they then undertook a simulated driving scenario where they needed to communicate and coordinate to carry out a transition of control from automated to manual driving. Cluster analysis was carried out on the transcripts of the verbal interaction between the two participants using Speech Act Theory (Searle et al., 1980). This theory purports that meaning can be conveyed from utterances in the form of an action, for example, a warning, question, advisement or statement. Speech classification types, adapted from the original theory for increased usability (Lampert et al., 2006), were applied to the simulated human-to-human interactions to identify how speech could be used to coordinate actions in human-autonomy interactions. The results provided evidence that suggested participants clustered their speech into eight marked interaction types specifically associated with either the automated system or driver role (4 types in each); additionally, salient driver-automation type pairing preferences could also be distinguished. Results also identified key information such as the frequency of dialogue turns, takeover time and, from those identified, the interaction types that participants were most likely to embody when handing control to one another. These rich and novel results can be used to inform research on HMI design, where naturally occurring human-to-human conversational speech data can be used to advance our understanding of how best to shape human autonomy interaction design and effectively address key concerns such as safety and trust.

The CHAT standard operating procedure presents a potential design template for use in HMI design. CHAT offers the standardisation of the procedural application of a checklist (which to the author's knowledge would serve as a novel HMI intervention approach). However, it also offers flexibility to be applied to the HMI design in a way that both fits with classic design guidelines, such as the ecological design principles (Rasmussen and Vicente, 1989), alongside innovative design ideas coming from the development of conversational user interfaces (e.g. Clark et al, 2024). A CHAT HMI has the potential to provide a novel and effective countermeasure to human performance problems relating to OOTL driver state induced by automated driving at level 3 automation. Additionally, as outlined earlier, this would also echo the effective tools used within the aviation domain and a potential method by which deliver

the CHAT SOP as a checklist of actions required of the driver (Trosterer et al., 2017).

6.4 Design Approach – Conversational Interfaces

Clark et al.'s (2024) research into naturally occurring human-human speech data highlights the potential use of conversational user interfaces within in-vehicle HMIs. This approach offers a flexible form of user interaction, which can deliver rich information in real time to drivers of automated vehicles. For this reason, Clark et al.'s (2024) research was used to inform the design of the CHAT HMI. The aim was to use the preferred 'automation' interaction types identified in the study to guide the design of 'how' information should be conveyed to the human driver to encourage human automation teaming and task focus during the transition of control. The CHAT SOP principles guided the design of 'what' information would be communicated to the driver.

Clark (2024) identified a number of clustered interaction types for both the role of the driver and the role of the automation. For the design of the CHAT HMI, it was decided that the 'supervisor' interaction type aligned best with the overarching goals of the CHAT approach and use of a standard operation procedure. This interaction type is concerned with both raising SA and directing the behaviour of others, aiming to ensure drivers have received critical pieces of information while also providing guidance on taking control. It primarily uses the speech act classification, 'Advisement', which attempts to guide behaviour through suggestions, commands, permission or prohibition (Lambert et al, 2006). Each speech turn using the supervisor interaction type is bounded, providing a definitive rather than open end to the turn by requesting that the driver takes a particular action towards completing the transition. Clark et al. (2024) found that the 'supervisor' automation interaction type paired well with 'driver' types who tended to acknowledge receipt of communication and self-readiness and raise their SA by inquiring about the environment. It is noted that in Clark et al.'s (2024) study these 'inquiries' were raised to the partner who was playing the role of the automated system. However, it is anticipated in this case, where the partner is the HMI display, that the driver will instead be prompted to seek these answers out themselves by engaging with the external environment.

Reflecting on key themes that emerged from the interviews conducted as part of the CHAT Driver Training study (see chapter five), the aim of using this type of interaction was to provide a consistent and incremental guide that would prompt drivers towards particular actions by bringing them into conscious processing. By communicating the CHAT SOP in this style, it aimed to direct driver behaviour and attention allocation in real-time to steadily build up and recalibrate driver SA and

workload across the transition period (Carsten and Martens, 2019). In essence, it aimed to encourage cooperation within the human-automation team, creating a concept of common ground (Klein et al., 2014) and shared control (Flemisch et al., 2012; Large et al., 2017a; Banks et al., 2019). This was focused, in particular, on the monitoring task and shared SA requirements, during the transition period and, importantly, ahead of the driver resuming full control of the driving task.

As highlighted in the literature (Endsley, 2017a; Carsten and Martens, 2019), it is important that drivers are given sufficient time and resources to raise their SA and workload on which takeover performance depends. Therefore, the transition timing within the HMI design was an important consideration. Clark et al. (2024) proffered that their findings demonstrated that users would likely want to engage in information-transfer during the transition of control. This would consequently lead to increased takeover time, which would need to be taken into consideration when designing the HMI. This was noted, alongside consideration of the upper bounds of 25.7 seconds reported in the seminal study on takeover times carried out by Eriksson and Stanton (2017a) and consistency with the minute takeover time provided in previous studies in this series (See Large et al., 2019; Shaw et al., 2020a; and Shaw et al, 2020b).

A subsequent design decision was made to keep the minute takeover time but split it into two equal parts. To reinforce the mental model of a planned transition of control and subsequently minimise automation surprises and reduce the psychological discomfort to drivers (Carsten and Martens, 2019), the first 30 seconds of the transition period was allocated to giving drivers time to switch their attention from any secondary task they were engaged with to getting ready to re-engage with the primary task of driving. To reflect the points made by Clark et al. and results from Eriksson and Stanton (2017), the later 30 seconds of the minute takeover time was allocated to turns relating to the key elements of the CHAT procedure and moving drivers into an 'on-the-loop' control state. Clark et al. (2024) found that the median number of speech turns were 3 per participant role (6 turns between driver and automated system). Therefore, three speech turns were allocated to the CHAT procedure of 'check' and 'assess', with a final instruction being the 'takeover' of manual control. A full description of the CHAT HMI is provided below.

6.5 CHAT HMI Design

The first 30 seconds of the CHAT HMI represented a standalone 'turn' of the 'conversation' and focused on task switching from any secondary, non-driving tasks, the driver was engaged in to focusing on preparing to resume the driving task. The visuals were designed to be clear and simple pictorial accompaniments to the speech used. Although there

was a visual representation, the design aimed for the voice communication to capture more of the driver's attention and convey enough communication so that it could be solely relied on in future drives and focus more of the driver's attention externally rather than internally. The aim of this first stage in the CHAT HMI was to direct the driver to disengage from their secondary tasks and physically prepare themselves for driving.

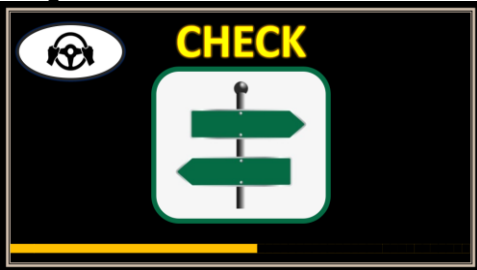
The first screen (see 'image 1' in Table 6-1) was designed to alert the driver to the takeover request and draw their attention to the HMI and voiced instruction. The written message changed from a green 'Automated Driving' to orange 'Prepare for Takeover' message. However, the centre automated driving mode symbol remained as a message to the driver that the ADS was still engaged. At the same time as the message changed there was an accompanying audio voice message that said, "Prepare for takeover". In previous studies within this series, the TOR HMI instruction was 'Prepare to Drive'. However, for both the CHAT and Basic HMI groups for this study the message was changed to 'Prepare for takeover'. The rationale for this was related to reflection that there was an underlying implicit temporal nature of the TOR instruction.

The interview analysis from the driver training study suggested that drivers, in particular those who had not undergone the CHAT behavioural training, were focused on taking over physical control of the driving task as soon as possible. The TOR HMI display screen was consistent for both HMI conditions. Therefore, it was important to reduce any potential confounding variables that may negatively bias the performance of drivers who were not exposed to CHAT during the takeover period. It was felt the 'Prepare to Drive' message in the original HMI design focused the mind of the driver on the physical action of driving rather than on the preparatory tasks needed ahead of the takeover. It was also felt that the implicit temporal order of activities and anticipated duration before driving shifted with the change from prepare *'to drive'* to prepare *'for takeover'*. With an implicit immediacy of being instructed to drive, i.e. carry out the physical action now, vs preparing for a takeover, i.e. carry out the required preparation activities to physically takeover the driving task in the future. As the aim for the HMI is to provide both groups adequate opportunity to take whatever actions they believe are required to prepare themselves for the takeover, it was felt that this change in language, albeit subtle, would add to the robustness when assessing the necessity for a behavioural intervention for planned control takeovers. Additionally, to again aim to reduce automation surprises and provide consistency, the symbols for automated driving (see image 1 in Table 6-1) and manual driving modes (see image 9 in Table 6-2) were used at key points within the HMI design.

Table 6-1 HMI Display Screens for the TOR and first 30 seconds of the transition period.

Image 1: 	Visual display to alert the driver to the takeover request. The centre logo for automated mode was used to indicate that automated mode was still engaged. The change in HMI display was accompanied by a voice message stating, <i>“Prepare for Takeover”</i>.
Image 2: 	Visual display to alert the driver to put away non driving related tasks in preparation for re-engaging with the driving task. The accompanying voice message stated, <i>“Put away non-driving activities ...”</i>
Image 3: 	Visual display to alert the driver to physically prepare themselves for resuming the driving task. The display image depicted a driver seated with their foot on the pedal, hands on the wheel and eyes looking out at the road ahead. The accompanying voice message stated, <i>“... and get into position for driving”</i> The display remained until the end of the first 30 seconds of the transition period.
Image 4: 	At the 30 second mark the visual display moved to alert the driver that manual driving mode would engage in 30 seconds, and they should begin takeover checks. Writing on the display read ‘TAKEOVER CHECKS’. The manual driving mode symbol with 30s in its centre was used to highlight that manual driving mode would engage in 30 seconds. The accompanying voice message stated, <i>“Begin takeover checks”</i>

Table 6-2 HMI Display Screens for the CHAT SOP & final 30 seconds before transition to manual driving mode.

Image 5: 	Check 1 consisted of the word 'CHECK' and a visual logo depicting a sign with directional arrows. The accompanying voice message stated, "Check for location and route signs". The symbol for manual mode moved to the top left corner of the display to remind drivers that manual driving mode would be engaged following takeover checks.
Image 6: 	Check 2 consisted of the word, 'CHECK' and a visual display depicting a bird's-eye view of vehicles on a road. Arrows point out in all directions from a centre vehicle towards the surrounding area. The accompanying voice message stated, "Check surrounding vehicles". The symbol for manual mode remained in the top left corner of the display.
Image 7: 	Display 3 in this series consisted of the word 'ASSESS', the manual mode symbol in the top left corner and the same bird's eye visual as the previous display. However, the arrows now protrude from the surrounding vehicles in the forwards direction and to the front left and/or right towards the adjacent lane. The accompanying voice message stated, "Assess and Monitor Traffic Movement".
Image 8: 	The final screen was a temporal instruction providing a visual countdown to the engagement of manual driving mode. It consisted of the words 'TAKEOVER IN' and the manual driving mode symbol with the numbers from 5 to 1 in the centre changing to coincide with the final 5 seconds of the takeover period. The accompanying audio was a combination of a voice message stating, "Takeover in" and 3 tonal pips over the final 3 second counts.
Image 9: 	At the point that manual driving mode engaged the HMI switched to the Manual Driving Mode screen.

Following the prepare for takeover screen ('Image 1'), 'image 2' was designed to communicate that non-driving tasks should no longer be used. A phone was chosen for the visual icon as this is a common secondary activity used during automated driving. The icon is also used in areas where phones are commonly not allowed, e.g. cinemas or theatres. Image 3 aimed to guide the driver to physically prepare themselves for the driving task. The visual image was designed to provide a pictorial cue of the activities the driver should focus on at this point. For example, proprioceptive realignment of their feet on the pedals, checking their seat and mirrors were in the correct position and placing their hands on the steering wheel. A pictorial reminder that NDRTs should not be used was also provided in this display. The final display was designed to provide a 30 second time check until manual driving mode engagement and to alert the driver that they should re-engage with the driver monitoring task and begin their takeover checks.

The second 30 seconds of the CHAT HMI represented 3 conversational 'turns' that were focused on guiding the driver to follow the CHAT SOP and move into an 'on-the-loop' of control driver state. The structure of the second half of the CHAT HMI design mirrored the behavioural training through its focus on the strategic, tactical and control subtasks of the driving task, with an emphasis on the spatiotemporal elements associated with situation monitoring for each of these levels. The aim was to highlight the goals of situation monitoring at each level of the driving task and instruct the driver through these goals as they transitioned from automated to manual system states. The HMI display used a combination of words, symbols, and pictures alongside audible, spoken 'advisements' and a countdown timer. The visual displays used in the HMI aimed to add meaning and structure to the transition procedure, in particular, in its initial use. The visuals were designed to provide a simple form of explanatory knowledge that would be efficiently understood by experienced drivers who were aware of the rationale behind the required actions from previous training and general driving experience. This use of imagery to provide additional meaning is particularly emphasised in images 6 and 7 (see Table 6-2), which used the same bird's eye view of the road as was used in the behavioural training (see Chapter four). These images were used to visibly convey the checks and assessment scans needed to the surrounding driving environment. In addition, the images aimed to prompt the driver to think about the potential movements of surrounding vehicles and, therefore, indirectly highlighted the consequences of not carrying out this procedure ahead of the takeover. As was highlighted in the driver training study interviews, these drivers may not recognise the need for these actions, in this order, within the context of a planned transition. These actions also tend to be performed automatically, below the level of consciousness. Therefore, it is important, especially in the early stages of using level 3 automation, to provide a prompt to show why

these actions are important in addition to the advisement to undertake them. Although it is recognised that some of these images were quite complex, they were designed so that they could be processed and understood at a global level and meaning could be drawn from them efficiently. Additionally, the overall concept behind the HMI design was for the visuals to support the audio and, therefore, for drivers to place less emphasis on the visual displays in future drives. With this in mind, the HMI could be offered as a standalone training tool prior to drivers undertaking their first 'live' drive in a vehicle fitted with these ADSs.

The CHAT SOP speech interactions were designed to sequentially raise driver situational awareness and workload through the use of 'advisements'. They were designed to prompt the driver's attention allocation and associated glance patterns beginning from: strategic goals, relating to spatial awareness and navigational cues; building to the tactical goals, planning future actions and required manoeuvres in relation to the positions and actions of surrounding road users; and finally control monitoring goals relating to controlling the basic vehicle operations required for resuming motor control of the vehicle following a transition to manual driving mode. The visuals were adapted from the CHAT behavioural training module and aimed to provide some context to the operating procedure by giving a bird's eye image of a driving scene. As with the other visual display screens for the CHAT HMI they were not designed to be the key focus of the driver's attention. Instead, the visual display screens aimed to provide additional visual guidance to benefit drivers in early use when they would be building initial knowledge and experience of using the HMI, whereas the speech interactions were intended to be the driver's primary focus of attention during continued use. In this way, the HMI aimed to facilitate guiding and directing driver's attention towards the external environment.

Overall, the use of conversational speech with the design of the CHAT HMI had a number of key aims:

1. Reflecting the research highlighting the importance of the cooperative requirements of the joint driving task in automated vehicles (e.g. Johnson et al., 2014; Klein et al, 2014a; Flemisch et al., 2012; Large et al., 2017a; Banks et al., 2019), the HMI design aimed to ensure system transparency and establish 'common ground' (Klein et al, 2014b; Eriksson and Stanton, 2017). The communication of conversational advisements from the vehicle HMI to the driver aimed to make it clear that both they and the automated system played a role in the shared responsibility of the driving task. This understanding and concept of 'shared control' is particularly important during the transition of control from the automated systems to the human driver and something that was not implicitly understood as highlighted by the driver interviews. Therefore, the HMI was designed to encourage drivers to develop a mental model of the automated vehicle as a joint cognitive system.

2. As per the HMI principles highlighted by Carsten and Martens (2019) the CHAT HMI design aimed to provide OPD (Johnson et al., 2014) to support appropriate behaviour and reduce automation surprises. The HMI provided the driver with clear and timely instructions related to the expectations of drivers regarding tasks such as situation monitoring required during the planned control transition. The timed phases of the transition built into the HMI were designed to provide a temporal guide and rhythm to the required actions from the driver, which clearly outlined the automation's intention to hand control to the driver (Lu et al., 2016). The phased format aimed to maintain the driver's psychological comfort by providing a gentle, sloping increase in responsibility and pace of action, which reinforced a mental model of a planned rather than emergency control transition. However, there was also a consistent and persistent focus on the driver's required preparation for the driving task to avoid any attentional shift from the driver back to engagement with their NDRT.
3. Following the ecological design principles of Rasmussen and Vicente (1989), the use of the CHAT SOP in the form of the final advisements, aimed to support the efficient, rapid and accurate recall of desired skill-based behaviours. The actions of the CHAT procedure communicated by the HMI system to the driver were designed to provide the efficient scheduling of situation monitoring tasks with the driver (a key procedural goal highlighted by Degani and Weiner, 1997) and encourage an 'on-the-loop' driver state prior to resuming control of the driving task.

6.6 Basic HMI

To assess the effectiveness of the CHAT HMI a baseline 'Basic HMI' was also created. This design focused on alerting the driver to the same 'Prepare for takeover' TOR and temporal count downs to manual mode engagement. The countdowns took the form of an incremental timer bar and a numerical 10 second count down for the final 10 seconds of the PCTP. The MDM symbol gave a visual alert for the driver to place their hands onto the steering wheel. The length of the PCT for the Basic HMI was also 60 seconds. Researchers studying control transitions have not agreed on an ideal lead time for control transitions (see Eriksson and Stanton, 2017a for a review). However, there is agreement within the literature that drivers should be given sufficient time to comprehend the current driving environment prior to resuming manual driving (e.g. Carstens and Marten, 2019; Gold et al., 2013; Endsley, 2017a). A report outlining the UK government's approach to L3 ADS states that drivers should be given a minimum of 10 seconds to undertake a control transition (Kinnear et al., 2020). Therefore, the baseline HMI was designed to provide an additional alert 10 seconds prior to manual driving mode engaged.

The design of the Basic HMI was deliberately simple to allow for an evaluation of the necessity for further training to be provided to experienced drivers or if providing sufficient notice of a pending change in control requirements was sufficient. Additionally, the simulated driving environment used for the evaluation study was designed to include a number of elements within the driving environment to draw the driver's attention during the drive and the transition period. This aimed to robustly test whether, given sufficient time, drivers would be able to use competence built from their experience and training in the manual driving context, to demonstrate both desired driver behaviour during the PCTP and the momentary capability to match the task demands following the takeover. Table 6-3 provides details of the Basic HMI audio-visual displays.

Table 6-3 Basic HMI audio-visual displays from the TOR to the resumption of Manual Driving mode (and disengagement from the automated systems)

	Visual display to alert the driver to the TOR. A timer bar provided an incremental countdown until it reached 10 seconds to manual driving mode engagement. The change in HMI display was accompanied by a voice message stating, <i>"Prepare for Takeover"</i>.
	At the 10 second point the text changed from 'PREPARE FOR TAKEOVER' to 'TAKEOVER IN' and the 'Manual Driving' mode symbol appeared containing a numerical countdown from numbers 10 to 1 in the centre. Each number changed to coincide with the final 10 seconds of the takeover period. The accompanying audio was a tonal pip at the 10 second point and 3 tonal pips over the final 3 second counts.
	Visual display to confirm the 'Manual Driving' mode system state. The appearance of this display coincided with count zero of the 60 second transition time and the simulated system disengagement of automated control.

6.7 Chapter Summary and Future Directions

The CHAT HMI design outlined in this chapter will be evaluated in a further simulator study in comparison to the Basic HMI described in the previous section. This study, which is outlined in the next chapter, aims to test, once again, the necessity of a driver intervention for planned control transitions. However, steps have been taken to improve the face

validity and ecological validity of the study design. It also aims to evaluate the effectiveness of the CHAT HMI both in terms of driver performance at the point of takeover, but also driver SA and performance following resumption of manual control. Finally, the study will test the retention of the effectiveness of the HMI tool over time through the use of multiple experimental drives. The hope is that CHAT will prove to be both an effective and flexible tool to support desired driver behaviour, both in the form of formal training interventions, and as a standardised and versatile check-list procedure, akin to those used by pilots, that can be applied to HMI design.

7. CHAT HMI Simulator Study

7.1 Introduction

This chapter describes the final driving simulator-based study carried out as part of the thesis. It provides a detailed justification and description of the experimental design and methodological approach, including the design of performance methods relating to measuring the situational awareness (SA) of drivers and linking that to an evaluation of driver performance.

7.2 Study Overview and Aim

This final driving simulator study focused on investigating the necessity and effectiveness of a technological design solution to guide and support driver behaviour during planned control takeovers (PCTs) in level three (L3) automation. The study aimed to evaluate the effectiveness of the CHAT HMI (see chapter six) at supporting desired driver behaviours in real-time during a planned control transition period (PCTP) from automated to manual driving modes (MDM). Effectiveness of the CHAT HMI was evaluated in comparison with a 'Basic HMI' across three simulated driving scenarios of varying task demand. The study also aimed to explore the necessity for the provision of additional guidance, or whether skills from manual driving experience and training are sufficient to safely perform the driving task in the context of L3 automated driving systems (ADS).

To date, most control transition research has focused on scenarios where the takeover request (TOR) has been issued to drivers due to safety and/or time-critical circumstances (Eriksson and Stanton, 2017b). Consequently, studies into driver performance have concentrated on investigating how quickly drivers can safely take over control of the driving task from ADS. Performance measures have typically included driver response times to takeover requests and their ability to takeover and perform the operational control of the vehicle (i.e. the steering and speed) at the point of transition. However, the bulk of research so far has neglected to focus on drivers' visual attention during and following control transitions (Pipcorn et al, 2024). Arguably, greater exploration into drivers' glance behaviour during control transitions would reveal further insight into the how drivers' visual attention when ADS are engaged impacts driver performance once manual driving is resumed. For example, Pipcorn et al. (2024) investigated the impact of automated driving (AD) on driver's visual attention before and after TORs on public roads. They compared the visual attention of drivers during manual driving with glance behaviour during AD before and after four takeover requests. Results showed that during AD, drivers reduced their attention to the forward road and increased duration of off-road glances compared to manual driving. However, critically, they found that levels

of visual attention towards the forward road did not return to levels observed during manual driving earlier than 15 seconds post transition. They concluded that drivers may takeover control before they are fully aware of the driving environment or potential threats. As such, this research highlights that it cannot be assumed that drivers are ready to respond to events shortly after the TOR and raises the importance of investigating driver's visual attention during control transitions and shortly after manual driving has been resumed.

According to Seppelt et al. (2017), the duration, timing and location of driver on-road glances dictates the visual information used by the driver to form expectations for how cues, such as the vehicle's position in relation to the movement of surrounding traffic, will change in the following moments. This moment-to-moment model of SA is used to make decisions and control responses in the management of the driving task. They highlight that vision science points to drivers' ability to acquire the 'gist' of a momentary scene rapidly. However, they stress that the amalgamation of dynamic scenes required for anticipatory hazard identification and response is more complex and requires a period of increased attentional allocation. They propose that the formation of appropriate expectations of events will inform the driver of when to time off-road glances in the moments leading up to precipitating events (i.e. the driver behaviour or environment state that initiates the near crash, crash or incident) and predict whether or not the driver makes ill-timed off-road glances near to safety critical incidents.

Seppelt et al. (2017) suggest that although current HMI evaluation guidelines focus on minimising off-road glance durations, this is only a part of the picture. They suggest that although their findings concur with this approach in the moments right before a safety critical incident, a broader perspective is necessary when looking further in advance of precipitating events. Their analysis of driver glance metrics during crashes and near crashes in a 100-car naturalistic driving study found that safety related behaviour is linked to the pattern of on and off-road glances used to support drivers' development of information about a situation. From their findings they propose that in advance of precipitating events, drivers are more able to anticipate hazards if they acquire information from the surrounding road scene. Accordingly, if driver's scanning of the road scene, including the surrounding areas is temporally interspaced and sustained, this will mitigate any reduction in awareness of the front roadway due to glances away from the road and have a cumulative positive effect on anticipatory control responses to hazardous events. Conversely, if the driver's attentional strategy does not allocate enough attention to the road between off-road glances this may have a cumulative effect of the opposite nature as subsequent glances may be made on the basis of incomplete SA.

It is proposed that the temporally punctuated guidance provided by the CHAT HMI will support appropriate attentional allocation and timely rebuilding of SA during the PCTP. It is anticipated that this will, in turn, result in more timely and consistent glance behaviour to the road and surrounding environment, resulting in a scanning pattern and attentional strategy that facilitates a positive cumulative effect on drivers' ability to make anticipatory control responses to hazardous events following the resumption of MDM. Additionally, it is predicted that the increase in SA and improved expectations of future events will result in lower levels of driver workload. This, in turn, will lead to improved driver psychological comfort and decision making, reducing the likelihood of a loss of control or driver errors that may lead to a near crash or crash. Additionally, it is proposed that the use of the CHAT HMI to communicate guidance to drivers via a checklist of required driver behaviours, provides, in real-time, the driver education and experience that could plug the gap in driver competence created with the new sub-task of PCTs. This, in turn, could mitigate the risk of limited driver capability being available beyond the point of MDM resumption.

This thesis purports that, without intervention, drivers do not engage in behaviours during the PCTP to appropriately recalibrate their SA and attention prior to resuming the driving task, even when given sufficient time. However, as per Fuller's (1999) Task Capability Interaction model (see section 2.9.1 chapter two), if the task demands are equal to or lower than the momentary available driver capability then the driver will likely perceive the task as manageable even where driver behaviour is less than optimal. Consequently, any negative impact of driver behaviour during the PCTP on driver performance post transition may be obscured if momentary driver capability is 'good enough' to cope with the demands of the driving task shortly after takeover. Therefore, to evaluate the effectiveness and necessity of the CHAT HMI, this simulation study manipulated the external task demands (see Figure 2-4; Fuller, 1999) the driver was exposed to in three different drives. Analysis of driver behaviour during and following a PCT, and performance outcomes relating to the management of driving task difficulty, allowed the level of available driver capability to be assessed against varying task demands of each drive. Subsequently, any resultant mismatch between available driver capability and task demands allowed the sufficiency of the HMI design used to be evaluated and an overall view to be formed on whether a behavioural intervention is necessary.

As such, the final study evaluated the effectiveness of the CHAT HMI compared with a Basic HMI in three simulated driving scenarios of varying task demand. The CHAT HMI intervention aimed to:

1. improve driver readiness by reducing shared attention with non-driving related devices (NDRTs), increasing driver behaviours to improve situation awareness, and transition to an 'on the loop' driver control state during the transition of control period.

2. improve driver capability by supporting appropriate calibration of SA, attention allocation and workload during the transition of control period to allow the driver to meet the different task demands of the driving task.
3. maintain driver behaviours referenced in (1.) consistently across the planned control transition period (PCTP) and over time across the three different drives.

It was predicted that drivers who experienced the Basic HMI will be more likely to have lower momentary driver capability following a transition to manual driving than those who experience the CHAT HMI. Thereby, these drivers would be more likely to experience a loss of control, resulting in evasive action or a crash when dealing with a high traffic demand scenario. It was also anticipated that there would be a difference in driver behaviour, in particular driver glance behaviour, during and following the PCTP between those drivers who experienced the Basic HMI in comparison to those who experienced the CHAT HMI. The expectation was that drivers who experienced the CHAT HMI would be more likely to demonstrate more glances towards the surrounding road environment than those who experienced the Basic HMI. Finally, where the task demands of the road were low or medium, it was expected that the momentary capability of all drivers, regardless of the HMI they experienced, would be greater than the external task demands of the road. Therefore, it was predicted that for these drives, there would be no statistically significant difference in the driver performance of vehicular control between the two driver HMI groups, as.

7.3 Methodology and Approach

The experiment was conducted in a medium fidelity, fixed based driving simulator at the University of Nottingham. The simulator comprises a right-hand drive, mark one Audi TT car, positioned within a curved screen allowing 270 degrees field of view to the front and sides of the vehicle. Three ceiling mounted, high-definition projectors are used to display the simulated driving scenario onto the screen with edge-blending and image warping to ensure a continuous, dynamic, unfolding scene is presented to participants. Two side mirror displays have been integrated into the original wing mirror frames and the rear-view mirror display is presented via a screen placed to the rear of the vehicle facing the vehicle's adjustable rear-view mirror placed in a standard position within the interior cabin. A Thrustmaster T500RS force feedback steering wheel and pedal set were integrated within the existing Audi primary controls and interior cabin. Dashboard vehicle performance data is presented on a 7-inch LCD screen, replacing the original Audi instrument cluster.

As with the previous studies, the driving simulator was modified to mimic the capabilities of a vehicle fitted with SAE L3 ADS. When

automation was engaged, drivers were able to relinquish all aspects of the dynamic driving task (DDT), which included both the physical primary operational control actions (steering, accelerating and braking) and monitoring tasks (object and event detection and response) to the vehicle's ADS. This capability had an operational design domain (ODD) that was geo-fenced to motorway driving in the context of the experimental scenario. In advance of entering and exiting the ODD, the ADS would present a sequence of audio-visual displays via the in-vehicle HMI. A 'wizard of oz' approach (Dahlback et al, 1993) was employed to remotely control the engagement and disengagement of the ADS (and accompanying HMI display sequences) by the researcher from the simulator control room without the driver being aware of this human involvement.

7.3.1 Experimental Design

The study used a 2 x 3 mixed factorial design with HMI design (CHAT vs Basic) as the between-subjects variable and task demand (low, medium and high) as the within-subjects variable (see Table 7-1). The within-subjects variable was used to compare differences in driver behaviour and performance over the three drives. It was also possible to evaluate how the drivers managed the difficulty resulting from the varying task demand and if there were any interaction effects from the type of HMI the drivers used.

Table 7-1 Experimental conditions

		Task Demand (within-subjects)		
		Low	Medium	High
HMI (Between-subjects)	CHAT	Condition 1	Condition 2	Condition 3
	Basic	Condition 4	Condition 5	Condition 6

The order in which the experimental drives were conducted was the same for all participants. As per the previous simulator study that evaluated the CHAT behavioural training vs operational training (see chapter five), it was predicted that drivers who experienced the Basic HMI would be less likely to demonstrate desirable driver behaviours during the PCTP than those who experienced the CHAT HMI. However, it was anticipated that, in accordance with Fuller's (1999) Task Capability Interaction model, any reduction in momentary driver capability would be more likely to impact performance where the task demand exceeded driver capability (e.g. during the high task demand scenario). As such, it was important that participants received an equal and adequate amount of experience interacting with the ADS and the HMI and conducting PCTs before undertaking the high task demand drive, which was the key focus of the study for performance comparisons. Participants were deliberately presented the medium task demand drive for the first study, to provide them with experience of particular events during the experimental scenario, for example, lane

changes and vehicles joining the motorway during the PCTP and the requirement for the participant to change lanes before taking the final motorway exit post MDM. The second drive was set as the low demand drive, to reduce the demand on participant's cognitive resources so that they could practice the PCT for a second time before tackling the third high task demand drive. As the HMI design was a between-subjects variable, order effects would not confound the results of HMI effectiveness between the two groups. Additionally, presenting the drives in a set order allowed an evaluation of the effectiveness of the HMI design over time.

7.3.2 Driving Environment and Vehicle Behaviour

A number of points from the literature and analysis from the driver training interviews (see chapter five) were used to shape the design of the simulated driving environment and the behavioural coding of the ego vehicle and surrounding traffic. The aim of the design was to improve both the ecological and face validity of the study, thereby increasing the robustness of the evaluation of both the effectiveness of the CHAT HMI and the necessity of a behavioural intervention for PCTs. The simulated driving environment was created using AV Simulation SCANeR software (<https://www.avsimulation.com/scaner/>). The use of this new simulation software allowed for additional fidelity and complexity to be introduced into the final driving environment and scenario, which was not possible with the STISIM software used for the driver training simulation study (chapter five).

Rule 254 of The Highway Code (DfT, 2023), highlights that the faster speeds used by traffic on motorways mean that drivers have less time to react. Consequently, the rule emphasises to drivers, the importance of using their mirrors earlier and look much further ahead than they would on other roads. As such, a motorway environment offers a useful setting to assess driver capability, in particular, driver skills such as SA and attention through the assessment of driving performance and glance behaviour. For example, higher levels of SA are associated with anticipatory hazard avoidance through increased attentional focus and levels of vigilance (Fuller, 2005; Seppelt et al., 2017), whereas drivers with low SA typically demonstrate more reactive behaviours when dealing with hazards (Endsley, 2018). Motorway driving also provides a realistic scenario relating to the anticipated operational design domain (ODD) for L3 ADS (e.g. Kinnear et al., 2021). In turn, this allows for the application of a predictable journey including a PCT e.g. a required exit from a motorway junction. This means that a consistent and controlled route structure can be used across multiple journeys, where the landmarks and route can become familiar to drivers. This route familiarity and predictability could also include the ODD boundaries i.e. when engagement of the ADS would become unavailable, and a TOR would be issued (Carsten and Martens, 2019; Trosterer, 2017). The predictability and reliability of the ADS would allow control requirements

to be consistently communicated to drivers in advance of the required exit junction. Therefore, this type of environment provides the ideal testbed for a simulated scenario that could evaluate the effectiveness of the CHAT HMI design. However, it also offers the opportunity to robustly test its necessity when the driver is provided with information they can use to predict when they will need to take control of the driving task; and sufficient time to regain SA and attention prior to resuming control from the ADS.

Interview analysis from the driver training study (see chapter five) highlighted that drivers wanted to experience how the ADS responded to other road users, thereby helping them form a mental model of how the vehicle's ADS operated and the capabilities and limitations of the system. This was particularly important in relation to the driver's understanding and expectations of their responsibilities, especially regarding PCTs and object and event detection and response (OEDR). Additionally, witnessing OEDR interaction would provide drivers with confirmation that the ADS was engaged and working well. Drivers reported that a lack of feedback during the driver training study influenced how comfortable they felt in automated mode and caused uncertainty regarding their level of responsibility in response to the movement of surrounding traffic. Carsten and Martens (2019) highlighted the importance of observability, predictability and directability (OPD) in relation to developing appropriate trust and psychological comfort of drivers. They emphasized that when the vehicle's systems are in control of the driving task, the driver will only know the vehicle's intended decisions and actions after it has begun to respond. Therefore, it is important for drivers' psychological comfort and reassurance that the vehicle's behaviour is explainable, in accordance with the predictable rules of the road, and demonstrates control and consistency.

Based on these considerations, the driving environment included a number of features for the ego vehicle ADS to demonstrate a conservative driving style, adherence to traffic rules and responses to other road users. As the ADS was programmed to engage on a quiet road prior to reaching the motorway, this gave an early opportunity to demonstrate the control and capability of the ADS as the ego vehicle needed to negotiate safe movement from the slip road onto the motorway. All traffic movement was controlled using triggers placed on the simulated roadway that would cause the vehicles to behave in particular ways, for example change lanes, speeds, or enter or leave the motorway. The ego vehicle was also coded to change lanes to pass slower-moving traffic. However, the behaviour of the ego vehicle was deliberately coded as conservative in nature, so that it provided the driver with more psychological comfort and amplified the control and reliability of the vehicle ADS. For example, the ego vehicle always adhered to the speed limit of the road and had appropriate lane discipline, pulling out only to overtake slow moving vehicles or to allow other road users entry onto the motorway at a junction. Lane changes

conducted by the ego vehicle were always preceded by vehicle indication and programmed to be smooth and steady, ensuring a large gap in the adjacent lane before commencing movement. Additionally, the ego vehicle's time headway tolerance was programmed to never go below 2 seconds. This provided the vehicle with a greater reaction time when responding to lead vehicle braking to allow the driver time to observe and understand vehicle behaviour and to give the impression of control. The presence of junctions within the motorway created predictability in relation to how the ego vehicle would behave (i.e. moving into the central lane if vehicles were joining the motorway). This, in turn, allowed slip roads to be used to set up scenarios following the PCT where the ego vehicle was in the central lane, to manipulate task demand at the point of takeover.

Finally, it is noted that drivers from the operational training group (see Chapter five) reported that the presence of salient changes in the driving environment would have drawn their attention and made them more alert. The behavioural evidence seen during the driver training study would suggest that drivers in the behavioural group demonstrated greater attention towards their external environment compared with the operational group, despite being presented with the same driving experience. However, to ensure validity and fairness in the evaluation and comparison of driver behaviour, the simulation must have good ecological and face validity. Therefore, for the current study, the driving environment and traffic movement was designed to be as close to a real-life scenario as possible (including the plausibility of the circumstances used to manipulate the task demands). Several landmarks (e.g. bridges and junctions) were included in the scenario to support driver attention, familiarity and predictability in relation to the timing of the TOR. Additionally, interactive traffic movement coded throughout the drive was designed to provide salient cues to draw the attention of drivers regardless of the HMI presented. These measures aimed to ensure the evaluation was as robust as possible within the limitations of a simulated environment.

The driving environment was designed to replicate a typical commuting journey where level 3 ADSs are predicted to be used. The environment began as a residential location, which was connected to a UK three-lane motorway via a slip road and entry junction and ended via an exit junction leaving the motorway. The environment was designed to replicate a geo-typical UK road network with layout, junctions, markings, and signage conforming to UK standards as far as possible, being guided by publicly available UK Government guidelines. Road names, numbers and locations were fictitious but used UK traffic sign guidelines in their design to adhere to driver's mental models of a UK motorway i.e. standardised colour, font, and lettering dimensions used for UK motorway signs (Department for Transport, 2018). There were no implied or intended associations with UK locations to avoid participants having a prior mental representation of an associated route that could

indirectly bias navigation or wayfinding decisions or judgements during the drives. The same road network was used for all drives, but the volume and behaviour of surrounding traffic differed between the drives to present varying scenario conditions and create the impression that they were different drives. With the exception of the test drive, all roads involved were populated with varying levels of traffic to reflect typical, dynamic road conditions.

As the AVSimulation SCANeR software used the metric system, all vehicle control data was presented and captured as kilometers per hour (kph). However, as the simulated environment was on a UK road network the vehicle dashboard speedometer and driving environment road signage referred figuratively to miles per hour (mph). In reality, this meant that participants received all speed related information in mph (and perceived that they were driving according to mph), but the vehicle control data was recorded at kph. For example, if the driver maintains the needle on the speedometer at 70mph, in accordance with motorway rules and signage, the vehicle control data will record the vehicle at travelling at 70 kph. All vehicles within the simulator were programmed to drive at the equivalent mph values in kph, therefore, the relative speeds were the same for all traffic in the simulated scenario. For example, a lead vehicle programmed to travel at 70 kph moved at the same speed as the ego vehicle being driven at a perceived speed of 70mph (or if the ego vehicle was in automated driving mode and the speedometer read that it was travelling at 70mph). In practical terms, all participants received the same experience, maintaining the validity of relative comparisons between the two groups. For analysis and reporting purposes, absolute values of kph, kilometres and meters are used.

7.3.2.1 Automation Engagement

To control the amount of time drivers spent in automation, the engagement and disengagement of the ADS was controlled by the vehicle rather than the driver. Drivers were notified during the study briefing that they would be told when automation was available and that this was linked to the ODD of the ADS. Participants were informed that the system would engage automatically, and they would be given a 3 second countdown via the in-vehicle HMI to notify them of the system state. Drivers were able to experience this feature during a test drive which was provided before they conducted the experimental drives. For pragmatic reasons related to control of the simulated scenario, it was necessary to start automation prior to the motorway, which theoretically is beyond the geo-fencing for this vehicle. However, this extended ODD scenario was explained to participants during the experimental briefing and as the researcher was in control of the transition this did not have

any bearing on performance or participant understanding of the experimental ODD boundaries for the ego vehicle.

During the driver training interviews (See section 6.4), drivers reported feeling more comfortable engaging automation when the road was quiet. Therefore, the point automation became available was designed to be a quiet straight, rural road to provide psychological comfort to the drivers. When the simulation's automated features had been triggered, the driver was given a degree of force feedback via alongside the audio-visual HMI display as confirmation. Further feedback was also provided once automation was engaged, as the steering wheel rotated in line with vehicle movements. An inactive trigger was placed onto the simulated roadway (only visible to the researcher) 650m before the entry slip road for junction 27 onto the motorway. This was used to cue the researcher to activate the timed PowerPoint for the HMI display (See Table 6-1, Table 6-2 and Table 6-3) and the keyboard trigger to put the ego vehicle into automated mode. In addition, by initiating automation during this phase of the simulated drive it ensured that all participants experienced the same traffic conditions throughout the motorway sequence (Figure 7-1).

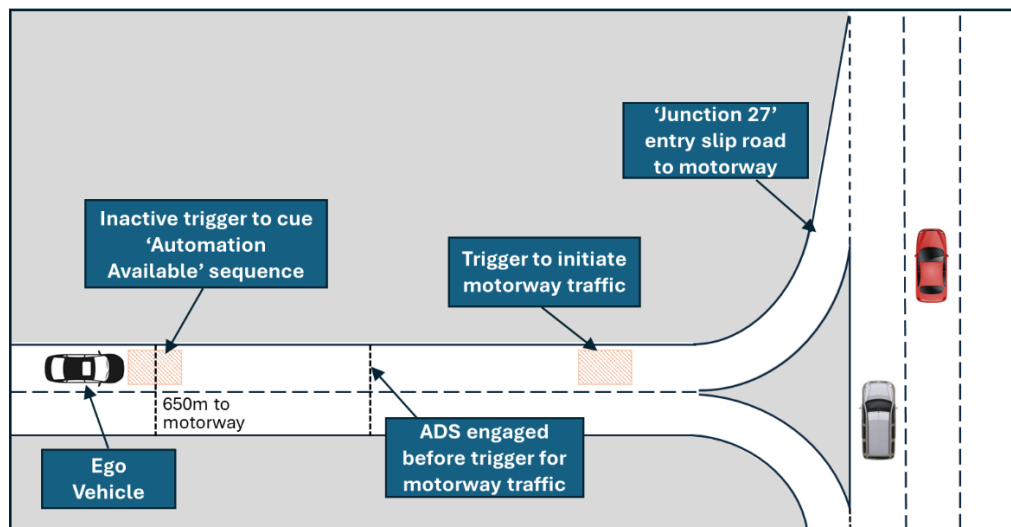


Figure 7-1 Schematic of automation engagement sequence ahead of the ego vehicle crossing the trigger to initiate motorway traffic.

7.3.2.2 Test drive

Prior to starting the experimental drive, all participants took part in a 'test drive' to experience and practise how to operate the vehicle. During the test drive, there were no other vehicles included in the simulation. The test drive began in a residential area with the car in manual driving mode. The participant driver pulled out of the parking spot and drove along a straight road towards the motorway. When the ego vehicle was 650m from the motorway slip road, the researcher triggered the HMI sequence for the ADS to engage and put the vehicle into automated driving mode (Figure 7-1). Once the vehicle was

established on the motorway (after approximately 1 minute), a TOR was issued, and the driver experienced the PCT HMI sequence and resumption of MDM.

7.3.2.3 Experimental drives

Three experimental drives were developed for the evaluation study. These were framed as commutes to a place of work or school and each lasted approximately 15-20 minutes. All experimental drives used the same route, beginning in MDM in a residential area. The residential road network was populated with typical features, e.g. trees and houses. It included a light level of traffic and a working set of traffic lights (Figure 7-2) to add to the immersive experience and provide landmarks to help familiarise the driver with the route.



Figure 7-2 Screenshot of the residential area with working traffic lights.

Once the sequence to automated driving mode was initiated, the ADS remained engaged for approximately 10 minutes. As with the automation engagement sequence, an inactive trigger was used to cue the researcher to instigate the PCT sequence on the HMI display at the same point in the journey for all 3 drives and all drivers. This trigger was placed 350m before the entry slip road of junction 29 and 2000m before the junction 29 exit slip road. Two landmarks were available to the driver to use as cues for the approaching TOR, as it was triggered just after the ego vehicle passed under a bridge and as the vehicle approached the junction 29 '1 mile' sign, visible to the left of the roadside. The TOR indicated to drivers that the ADS was reaching the boundary of the ODD (i.e. taking the next motorway exit) and the driver was required to undertake a planned control transition. The HMI display sequence began with the TOR alert and the driver was given 60 seconds notice before the transition to MDM. MDM began 800m to junction 29 entry slip road and shortly after the ego vehicle passed the junction 29 '½ mile' sign, visible to the left of the roadside. After MDM had engaged, drivers were required to leave the motorway at the exit slip road for junction 29. Once the driver had taken the slip road that marked the end of the journey.

Each experimental drive differed in the scenario presented to the driver during and following the PCT. A combination of features from the driving environment, based on Fuller's (1999) definition of task demand, were used to create a low, medium and hard task demand scenario. For example, traffic density was manipulated using different numbers of vehicles joining the motorway and the ego vehicle was positioned in the central lane for two of the scenarios, which subsequently required additional actions from the driver to successfully exit the motorway e.g. change lanes. The drivers all undertook the drives in the same order: medium demand, followed by low demand, and then hard task demand. Details of the scenarios following the issues of the TOR are presented below:

7.3.2.3.1 Scenario 1: Medium task demand

A single, slow-moving truck joined the motorway at the entry slip road for junction 29 and remains in lane 1 (the slow lane). The ego vehicle moves from lane 1 to lane 2 (the central lane) to give space for the truck to join the motorway. The ego vehicle (70kph) remains in lane 2 for the remainder of the PCTP as it proceeds to overtake the slower moving truck (60kph).

MDM engages as the front bumper of the ego vehicle in lane 2 is in line with the rear-bumper of the truck in lane 1. There are no vehicles directly to the front or rear of ego vehicle. There is one vehicle in front of the truck in lane 1, this vehicle is moving at 70kph and has an expanding time-headway (THW) of a minimum 4s from the truck. Therefore, in this scenario the driver has the option to continue to pass the truck and move into lane 1 ahead of the truck or brake and move into lane 1 behind the truck. Both were viable options within the speed limits and distance to the junction exit. A scenario schematic is presented in Figure 7-3.

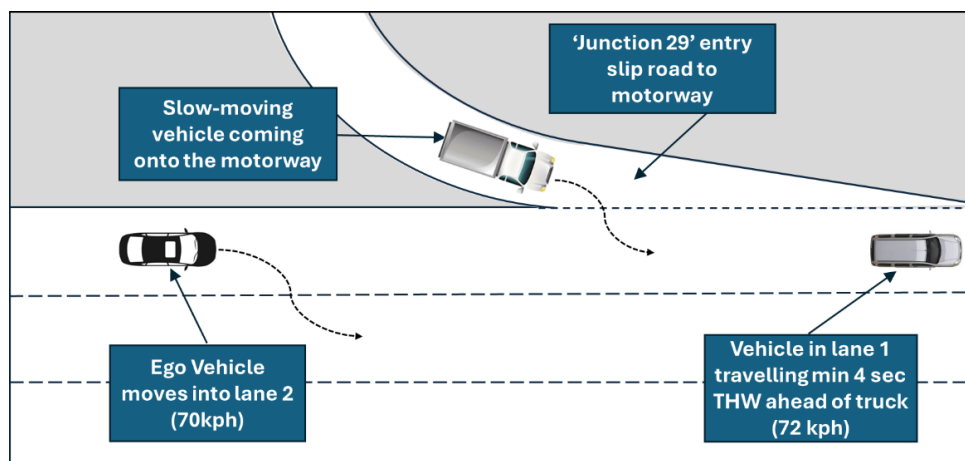


Figure 7-3 Schematic of medium task demand scenario as the ego vehicle approaches the entry slip road of junction 29 during the PCTP.



Figure 7-4 image of road centre from the viewpoint of the driver of the ego vehicle at the point of MDM engagement.

7.3.2.3.2 Scenario 2: Low task demand

No vehicles joined the motorway at junction 29 and the ego vehicle is positioned in lane 1. There is one vehicle travelling in front of the ego vehicle at the same speed and a THW of 2s. Following MDM engagement there are no vehicles to the rear of the ego vehicle, but there are vehicles ahead in lanes 2 and 3 (the vehicle in lane 3 overtook the ego vehicle during the PCTP). As the lead vehicle is travelling at the speed limit and at an appropriate time headway this provides a vehicle that the driver can follow until it reaches the exit for junction 29. A scenario schematic is shown in Figure 7-5.

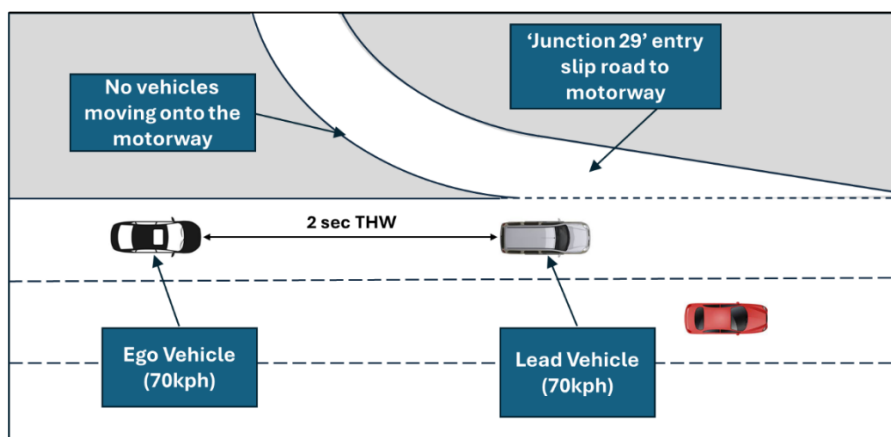


Figure 7-5 Schematic of low task demand scenario during the PCTP. Ego and lead vehicle remain at the same speed and THW for full transition period.

7.3.2.3.3 Scenario 3: High task demand

This scenario was adapted from a simulated drive used to investigate out of the loop gaze behaviour patterns of drivers conducting a critical control transition in a level 3 vehicle (Schnebelen et al., 2020). Four vehicles join the motorway at the slip road for junction 29. The ego vehicle moves into lane 2 to provide space for the vehicles to join the motorway. A white truck joins 75m ahead of the ego vehicle travelling at

67kph and remains in lane 1. A vehicle (lead 1) joins the motorway ahead of the ego vehicle and moves into lane 2 at 45m distance headway from the ego vehicle. The next two vehicles (lead 2 and lead 3) join the motorway behind the ego vehicle. Lead 2 moves in lane 2 100m behind the ego vehicle and lead 3 moves into lane 3. Both lead 2 and lead 3 begin to accelerate. Scenario schematics are shown in Figure 7-6 and Figure 7-8 and images from the left and rear view mirrors are shown in Figure 7-7.

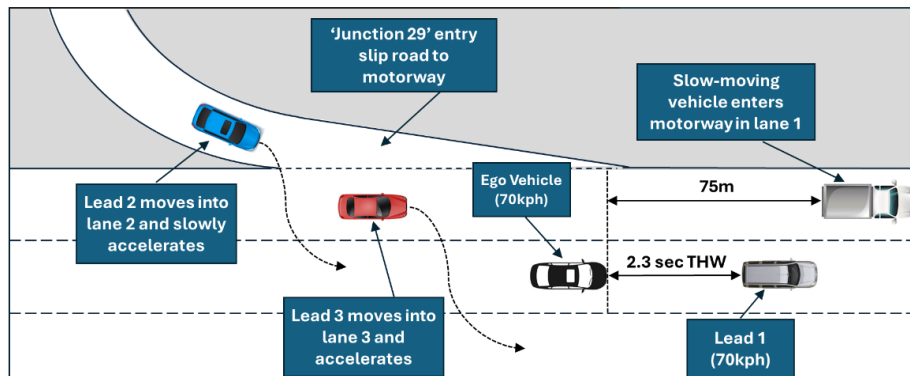


Figure 7-6 Schematic of vehicle positions during the PCTP as the ego vehicle moves passed the entry slip road for junction 29

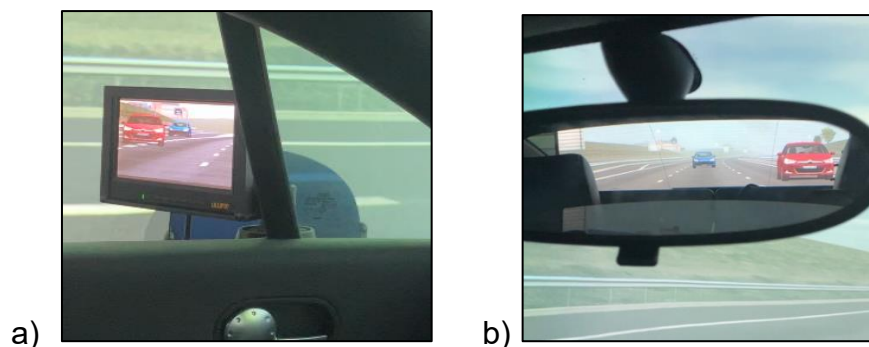


Figure 7-7 Example images taken from the driver seat, when the ego vehicle is in lane 2. Image a) shows the left-hand mirror view of lead 2 and 3 vehicles moving onto the motorway from the entry slip road at junction 29 (left). Image b) shows the rear-view mirror image of lead 3 in lane 3 and lead 2 in lane 2.

When lead 2 is 45m from the ego vehicle and 5s prior to manual driving mode engagement, it moves from lane 2 into lane 1 and continues to slowly accelerate. At the point that manual driving mode engages, lead 2 is 40m to the rear of the ego vehicle in lane 1.

Shortly after manual driving mode has resumed (approximately 3s) the lead 1 vehicle, which has started to overtake the white truck, brakes hard suddenly for 2.5s. At this point, the lead 2 vehicle is 30m to the rear of the ego vehicle in lane 1 and lead 3 is proceeding to overtake the ego vehicle in lane 3. There are no vehicles to the rear of the ego vehicle. Immediately following the hard brake action, lead 1 indicates and moves into lane 1 behind the white truck.

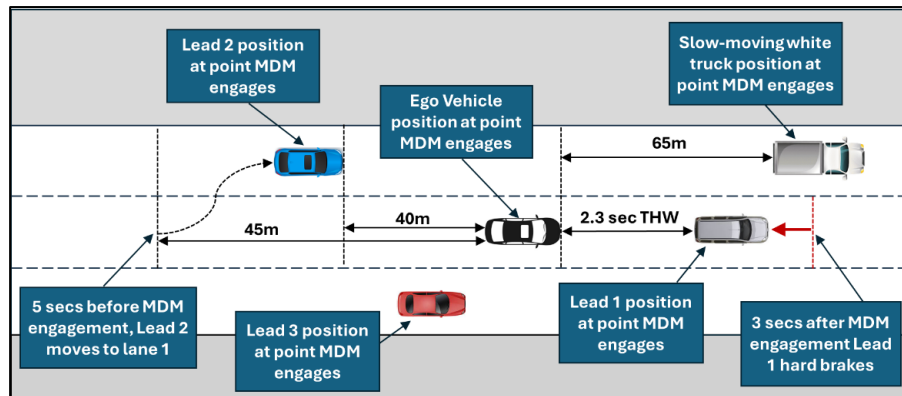


Figure 7-8 Schematic of vehicle positions at the point of MDM engagement and the movement and actions of Lead 2 (before MDM) and Lead 1 (after MDM)

In this scenario, the TOR for the control transition is planned and issued due to the approaching ODD boundary (as per scenarios 1 and 2). However, after the PCT has been completed, the driver is required to respond to the critical safety incident under manual driving conditions, demonstrating the sudden, urgent and unpredictable nature of task demand in the dynamic driving environment (Teh et al, 2014). Once the driver has responded to the actions of lead 1, they then need to decide when and how to get into position to take the upcoming junction exit. There are vehicles in lane 1 and lane 3 and the white truck is too far ahead for the driver to overtake that vehicle and make the exit without significantly breaking the speed limit. It is possible to move into lane 1 ahead of lead 2 or lead 1 who are both now positioned in the slow lane. However, the gaps between both the truck and lead 1, and lead 1 and 2 are small. As there are no vehicles behind the ego vehicle, the safest option would be to gently slow down and pull into lane 1 behind lead 2.

The scenario is deliberately set up to test the driver's SA, specifically in relation to tracking the movements of the lead 2 vehicle. To safely respond to these deliberately designed fluctuating task demands, drivers will need to be able to select, perceive and interpret relevant objects and elements from the traffic situation. This will allow them to carry out effective decision making and action response (Teh et al. 2014). It is predicated that drivers who have not raised their attention and SA in a timely manner during the PCTP will be unaware of the movements and location of lead 2, which is positioned (depending on how hard the driver brakes in response to lead 1) in the blind spot of the ego vehicle. It is anticipated that these drivers will be less likely to possess the ability to effectively and proactively adapt to the dynamic task demands of the driving environment and more likely to demonstrate reactive behaviours to the hazardous situation (Fuller, 2000; Endsley, 2017a; 2018), for example, changing lanes at an inappropriate point, without appropriate attention allocation and SA that could result in a collision or require evasive action. Additionally, a spike in workload due to inconsistent, insufficient or inappropriate glance

behaviour prior to the transition of control, may lead to poor driver decision making or deleterious effects such as cognitive tunnelling (Endsley, 2005) or inattentional blindness (Herslund and Jørgensen, 2003).

7.3.3 Data Structure and Performance Measures

This section outlines the data structure and performance measures used for the final study. The study produced a vast amount of data, therefore, to ensure analysis focused on the research goals, the main areas of interest were directed towards the following:

1. Driver actions during the PCTP. For example, the discharging of NDRTs, first glance at the roadway, hands on the steering wheel etc.
2. Driver glance behaviour during and following the PCTP. This includes attentional allocation towards pre-determined areas of interest (AOIs) (see Table 7-3) during specified event blocks (Table 7-2).
3. Driver SA and management of the driving task in relation to varying task demand. This predominantly focuses on the high task demand scenario and quantitative and qualitative analysis of driver attention (e.g. glance behaviour) and decision making (e.g. evidence of anticipatory/reactionary control responses or driver error).

7.3.3.1 Event blocks

To structure the results, data were selected and cut into the following event blocks for analysis (Table 7-2).

Table 7-2 Event blocks for analysis

Event block	Description/Definition
Total Planned Control Transition Period (PCTP)	This is the total 60 second time period between the issue of TOR and engagement of manual driving mode (MDM)/disengagement of the automated driving system (ADS)
PCTP T- periods (T1 to T6)	The PCTP was split into 6 x 10 second blocks with the first 10 seconds (T1) starting at the issue of the TOR and each 10 second block following until MDM engagement.
T-7	The is a 10 second block immediately following MDM engagement.
Total Lane Change	The starting point for a lane change event was when a vehicle made a continuous movement towards a lane boundary (i.e. with not reverse movement) and ended when the vehicle was fully present in the target lane with the last movement before the vehicle began to move in the opposite direction (Lin and Bao, 2019).
Lane change T- periods (T-10, LCB and LCA)	Behaviour relating to the lane change was broken down into the 10 second block before the lane change began (T-10); the time period before the vehicle crossed the lane boundary (LCB) and the time period after the vehicle crossed the lane boundary (LCA).

These event blocks allowed for driver behaviour to be segmented and analysed both individually within each block (and sequentially within blocks with consecutive temporal segments) and relatively between different event blocks. The aim was to provide additional context to driver behaviour and to investigate the transactional relationships between that behaviour (in relation to driver capability) and the events and features within the simulation (in relation to external task demand). For example, driver behavioural patterns in relation to the presence of external stimuli, such as vehicles moving onto the motorway from the slip road or the HMI display sequence (in particular the sequence of guided actions during the CHAT SOP); and relative patterns between driver glance behaviour during the PCTP and glance behaviour during the lane change.

7.3.3.2 Vehicle control and driver behavioural data:

Vehicle control data for the ego vehicle and key vehicles involved in the scenarios during and following the PCTP were captured by the AVSimulation SCANeR software at a sampling rate of 50ms (20 Hz). Key data metrics for each experimental drive were extracted and collated for analysis (for full details see Appendix F). Split-view cameras were used to capture video throughout the study to enable video analysis of driver behavioural data (Figure 7-9). Cameras were positioned so that the internal and external environment could be recorded and observed. Behaviour Observation Research Interactive Software (BORIS) (Version 7.4.7) (Friard and Gamba, 2016) was used to code salient driver behaviours from the recorded videos of the experimental drives.



Figure 7-9 Video split-screen image of driver engaging in a NDRT (reading a book) during ADM after switching on the in-cab reading light.

Frame-by-frame coding of behavioural observations was conducted from the split-screen video recordings of the experimental drive. Coded observations were time-stamped with onset and offset times for duration calculations. Key events such as the TOR and MDM engagement and lane change movements of the ego and surrounding vehicles were also coded within the observations. This allowed driver actions to be analysed temporally and sequentially, to produce metrics such as time (s) from the TOR to driver hands on wheel. This also provided anchor points to facilitate data triangulation, using the BORIS coded data, video recordings and the simulation data, to align vehicle control movements with driver behaviour for analysis. Coded behavioural data were extracted and grouped according to the event blocks outlined in Table 7-2 for analysis.

The ethogram created for previous studies in this thesis (see Chapter 3 and Chapter 5) was adapted and used for this final study. Behavioural indicators included salient actions such as discharging NDRTs (complete disengagement with no further reengagement), placing hands on the steering wheel, onset and offset of indicators etc. The predominant data extracted from the video recordings was driver glance behaviour. Specific AOIs were identified and included in the ethogram and grouped together for analysis (Table 7-3). An example of driver glances directed towards the rear-view, left-wing and right-wing mirrors is provided in Figure 7-10.

Table 7-3 Glance categories and areas of interest

Glance Category	Areas of interest included in the glance category
On Road - roadway	Road front centre
	Road Front left
	Road front right
On Road - periphery/rear	Passenger window (Blind spot)
	Driver window
	Rear view mirror
	Left mirror
	Right mirror
Off road - internal	HMI
	Dash
	NDRT
	Feet/pedals
	Passenger seat
	Other internal (all gazes directed outside the above areas)

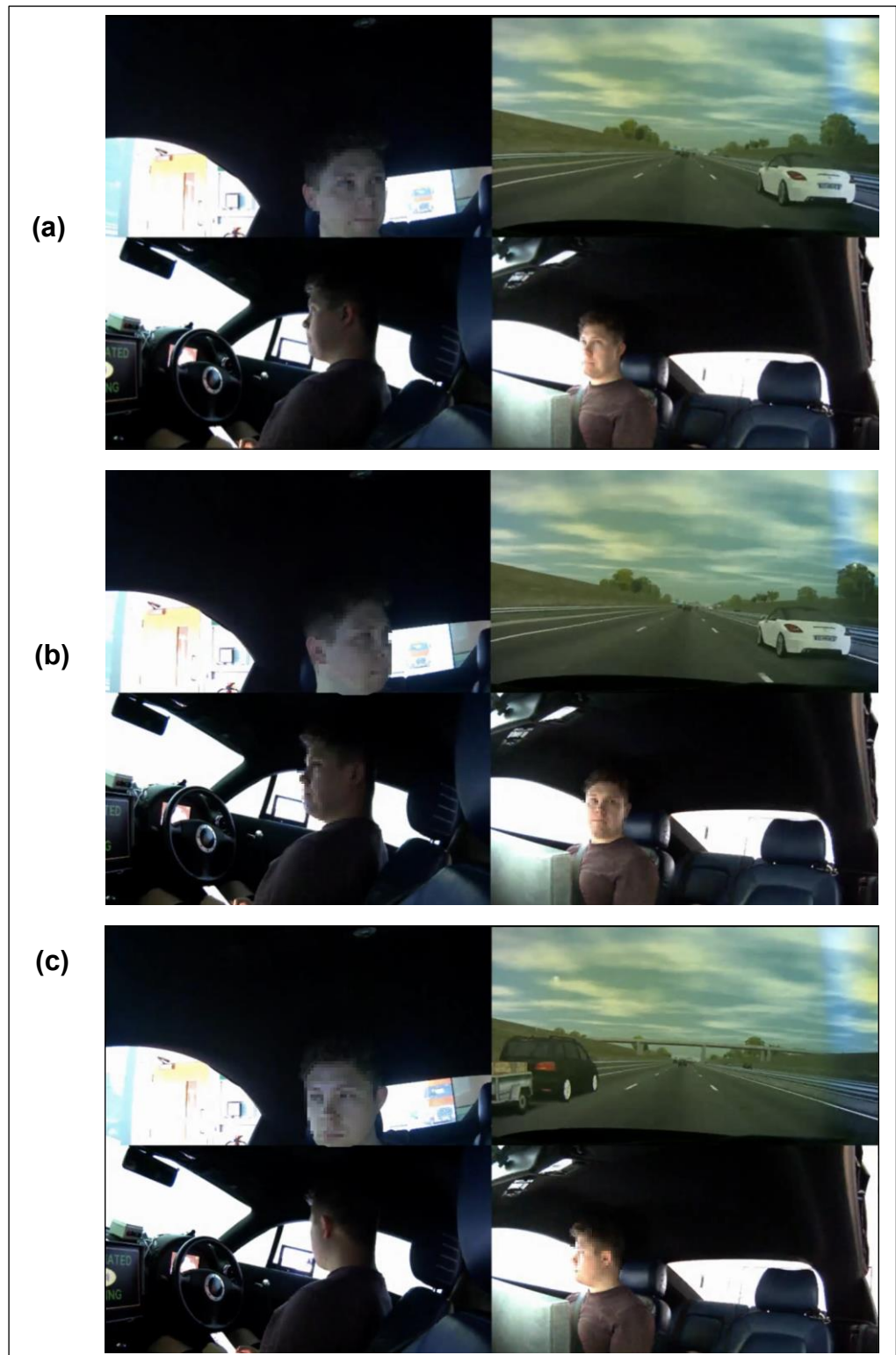


Figure 7-10 Images of drive glancing at (a) rear-view (b) right-wing (c) left-wing mirrors

7.3.3.3 Performance metrics

Vehicle control data provides key driving performance metrics associated with takeover performance (see Ostlund et al. (2004), for a detailed summary). For example, longitudinal and lateral acceleration,

standard deviation of lane position and steering wheel angle (SWA) aimed to identify the smoothness of the resumption of control and lane change manoeuvre. However, the key utility of these data for the current study related to the exploration of visual attention. These data could be used to precisely identify where the ego vehicle was in relation to the surrounding vehicles. This meant that control responses could be analysed both in terms of the position and proximity of the ego vehicle to surrounding vehicles during key manoeuvres, e.g. control inputs that indicate a rapid, evasive manoeuvre; lane exceedances; or time to collision calculations etc. during a lane change.

Additionally, visual observations of driver behaviour (e.g. glances to AOIs, indication onset, startle responses etc.) could be precisely cross referenced with the movements of the vehicles within each scenario to provide quantitative and qualitative data context to analyse the performance of driver SA and decision making. For example, during the high task demand scenario, it was possible to link driver's awareness of the movements of the surrounding vehicles, in particular, the lead 2 vehicle, during the PCTP with resultant control responses to the hazardous situation and required lane change following MDM engagement. This allowed robust inferences to be made regarding driver capability e.g. evidence of anticipatory or reactionary/loss of control inputs. Vehicle control data was cross referenced with driver behavioural data and information on the visual field of view (fov) of vehicles in the rear and side mirrors, which allowed analysis of the number of glances to the left and rear mirror when the lead 2 vehicle was visible and moving across the lane boundary into lane 1. These data coupled with elements such as follow-on or repeated glance data, and the timing and movement of the subsequent lane change allowed for analysis and an inference of SA performance to be made.

7.3.3.3.1 Subjective Ratings

Prior to taking part in the test and experimental drives, participants completed the Total Trust in Automation Questionnaire (TTAQ) (Gold et al., 2015b). Following each of the experimental drives, drivers were provided with a set of subjective questionnaires to complete. These included a further completion of the TTAQ; the Situational Awareness Rating technique (SART) (Taylor, 2017) and the NASA-TLX workload questionnaire (Hart and Staveland, 1988). Drivers were explicitly directed to answer these questions specifically in relation to the PCTP and manual driving that following the transition of control.

7.3.3.3.2 Driver actions during PCTP

Salient driver behaviours were captured from the recorded videos during the experimental drives to analyse how drivers physically prepared for driving following a TOR (e.g. time to first glance to

roadway, hands on wheel etc.). A particular focus was placed on the drivers' shared attention between the driving task and their NDRTs. A key part of the methodology for this study (and the series of studies within this thesis), was the use of the driver's own NDRT rather than a prescribed secondary task. This enabled analysis of drivers' willingness/motivation to cease using their NDRT and turn their attention towards the driving following a TOR. Shared attention was defined as the time spent between the onset of the TOR to the point the driver fully discharges their NDRT (i.e. putting the NDRT away and not returning to it).

7.3.3.3 Glance Behaviour

To get an understanding of the driver's attention allocation, driver glance behaviour towards pre-defined AOIs (see Table 7-3) was analysed. On and off-road glance behaviour provide important metrics to understand the attentional allocation and sampling rate of drivers. Therefore, analysis of these key performance metrics will give an indication of where, and for how long the driver focused their attention. Subsequently, inferences can be drawn on whether the driver has detected key cues within the driving environment that may signal a potential hazard, facilitating anticipatory response behaviour (Seppelt et al., 2017).

Glance frequency (glance per minute) and duration (the coded onset time of the current AOI glance until the coded onset time of the follow-on glance to a different AOI) metrics were calculated from the BORIS coded visual data output. Data were interrogated by exploring the variability in driver glance behaviour between the two HMI groups, during and following the PCTP. Analysis was carried out on each of the event blocks detailed in Table 7-2. Key events that occurred within the experimental drives are used to add context to the results. For example, during T-periods 2 (T2) (10-20s) and T3 (20-30s) within the PCTP, the ego vehicle was passing the entry slip road for junction 29. For two of the drives (1 and 3), there were vehicles present on the slip road and the ego vehicle pulled into lane 2 to give them space to move onto the motorway. This accompanying context can help to provide additional explanation for observed driver behaviour.

7.3.3.4 Driver capability and control response

Analysis of driver capability and control response was of particular importance in the evaluation and comparison of driver behaviour during the high task demand scenario. Investigations aimed to evaluate if:

1. drivers in the CHAT HMI group had better SA compared with the Basic HMI group.

2. drivers in the Basic HMI group were more likely to experience loss of control leading to near crash/crash involvement (see) compared with the CHAT HMI group.

To complete these evaluations, it was necessary to investigate the best way of measuring SA in the use of a complex system such as an automated vehicle. The importance and complexity in measuring SA in a reliable and valid way was reported by Salmon et al. (2009), alongside a variety of approaches available options with their associated pros and cons. It was important to provide a view of the driver's subjective assessments of their SA as well as quantitative and qualitative analysis of SA from data captured during the simulated drives. Therefore, this study employed a self-rating technique and objective measures from the experimental drive to measure and evaluate driver SA. A battery of tests was selected to increase the reliability and validity of the analysis and combat the doubts involved in the validity of subjective measures of SA (Salmon et al., 2009).

As the study was focused on providing as immersive experience as possible, SART was selected as the most appropriate subjective assessment method for the study, despite its lower validity (Salmon et al., 2009). Process indices from the experimental data were selected to analyse driver SA during the PCTs and lane change manoeuvres. Glance frequency, duration and percentage allocation to AOIs data, with reference to time periods within the PCTP and cues from the external driving environment, were used to measure and evaluate driver SA during the PCT for each of the experimental drives. Additionally, for drive 3, a combination of driver glance behaviour during the PCTP and the lane change was analysed, alongside vehicular control input data from the ego vehicle and surrounding traffic to provide context. The complete data 'picture' was used to assess the level of driver SA based on driver attentional allocation precipitating the lane change and driver performance during the lane change task itself.

Additionally, it was also necessary to decide how best to evaluate driver capability in relation to a loss of control and, specifically, a near crash. Seppelt et al. (2017), define a near crash as any situation that requires a rapid, evasive manoeuvre by the ego vehicle or any other vehicle, to avoid a crash. They go onto define a rapid, evasive manoeuvre as any form of control input that approaches the limits of the vehicle's capabilities, for example, steering, braking, accelerating or any other control response that has the same effect. Therefore, evidence of control inputs that fall into this category will allow a clear separation of

driver performance into 'loss of control = yes' and 'loss of control' = no, alongside supporting evidence from driver lane change performance metrics and glance analysis.

Critical jerk has been shown to be a successful performance measure for detecting near crashes and more reliable than other measures such as longitudinal acceleration (Bagdadi, 2013). Therefore, it is a good indicator of near crash involvement and source of information to use in the evaluation of the driver behaviour in each of the HMI groups. Critical jerk or 'abrupt onsets of brakes' is defined as the occurrence of a deceleration change induced by braking (or 'jerk') higher than a set threshold level. The presence of abrupt onsets of brakes at the beginning of a deceleration, can indicate the occurrence of a critical situation (Ostlund et al, 2004). SCANeR simulation braking and acceleration data from the ego vehicle was used to calculate jerk ($j = (\Delta a(t)/\Delta t)$, where a is acceleration, and t is time. The number of abrupt onset of brakes (at a threshold level jerk $> 10 \text{ m/s}^3$) was recorded for each driver during a braking response in relation to the lane change. If the driver made no control input to the brake, this was recorded as 0, so that all participants were included in the analysis.

Driver glance behaviour leading up to the lane change (in particular during a just after the PCTP) was used as evidence of the precipitating factor of the driver. This, according to Seppelt et al. (2017) is defined as the driver behaviour or state of the environment that initiates the crash, near crash or incident and the subsequent sequence of actions that results in the incident, crash or near crash. Analysis was carried out on individual driver performance, looking at the relationship between the driver's attention allocation (i.e. glance behaviour) and their actions during the lane change, in particular with reference to the position of the lead 2 vehicle. This was then used to assess whether near crash involvement took place and categorise the drivers accordingly.

7.3.3.4 Statistical Testing

Testing for statistical significance was conducted using SPSS software. The primary aim of this research was to evaluate the effectiveness of the CHAT HMI design on driver behaviour in comparison to the Basic HMI. Therefore, independent samples t-tests were carried out on all performance measures to compare the means of the drivers in each HMI condition. Where data sets had a non-normal distribution, non-parametric Mann-Whitney U tests were used to make statistical comparisons between groups.

A secondary aim was to evaluate the effectiveness of the CHAT HMI in the context of different task demands over the course of multiple drives.

Therefore, a two-way repeated measures ANOVA (mixed design) was employed to evaluate the effects of the drive task demand (within-subjects independent variable) on performance measures of the drivers in each HMI condition (between-subjects independent variable). Using a two-way repeated mixed design ANOVA also allowed the evaluation of between-subjects effects for each of the drives. Within-subjects effects relating to recognised rating scales, employed to collect subjective views from drivers on their own behaviour and the demands of the task, were also used to evaluate the face validity of the scenario designs. Appropriate corrections were applied to tests of within-subjects effects. According to Field (2018), Mauchly's test of sphericity has too little power for small samples. Therefore, for all within-subjects main effects a correction was applied based on the value of the Greenhouse Geisser Epsilon. If $\epsilon < .75$ the Greenhouse-Geisser correction was applied, otherwise Huynh-Feldt correction was applied. Additionally, as the drive condition had three levels, Bonferroni corrections were applied to all pairwise comparisons. Where data sets had non-normal distributions, non-parametric Friedman tests were employed to investigate statistical differences within drives or T-period per HMI group. Post hoc Wilcoxon signed rank tests with Bonferroni correction applied to the appropriate level were conducted to investigate paired comparisons where Friedman test results were statistically significant.

Finally, it was anticipated that in drive 3, drivers who had the Basic HMI would have lower SA (e.g. not be aware of the lead 2 vehicle) and be more likely to make driver errors or be involved in a near crash incident compared with the CHAT HMI group. Chi-squared and Fisher's exact tests were used to analyse the categorical data resulting from the classification of this behaviour. It was employed to examine the significance of the association between the two kinds of classification. For example, whether drivers made mirror glances when the lead 2 vehicle was changing lanes and in the field of view of the mirror or made no glances during this time.

7.3.4 Participants

Twenty-eight participants were recruited to take part in the study. All participants were considered experienced drivers, i.e. held a valid UK driver's licence, had a minimum of 3 years driving experience, including experience of motorway driving, and drove on regular basis (at least 2 days per week). Participants were primarily employees or postgraduate students from the University of Nottingham, recruited using advertisements placed across the campus or circulated via email. Four participants were unable to complete the study due to simulator sickness. Therefore, 24 participants completed the full study (14 male, 10 female; mean age: 37 years; range: 23-64 years; mean years of driving experience: 16; range weekly miles driven: 15 - 300). Participants were matched as closely as was practical for age, gender

and years driving experience between the 2 HMI conditions (See Table 7-4).

Table 7-4 Participant demographics by experimental condition.

	CHAT HMI	Basic HMI
N	12	12
Identified Male	7	7
Identified Female	5	5
Mean Age (years)	35	38
Mean years driving	15	17

Participants attended the study for approximately 2 hours and were reimbursed with £20 in shopping vouchers as compensation for taking part. The Faculty of Engineering's Ethics Review Committee approved the study procedure.

7.3.5 Human Machine Interface:

An in-vehicle HMI was installed in the centre console of the vehicle. This delivered audio-visual feedback and alerts throughout the drive via interactive, time-controlled PowerPoint presentations. Feedback from the HMI included whether ADS was available, the current driving mode ('Automated Driving' or 'Manual Driving') and notifications and guidance relating to a TOR and PCT. The delivery of HMI display screens and associated engagement and disengagement of ADS were remotely controlled by the researcher from the simulator control room. During the study, drivers were either presented with the CHAT HMI design or the basic HMI design. Both groups were given the same amount of time (60 seconds) between the takeover request (TOR) and MDM engagement. Each group was presented with the same design of HMI (CHAT or Basic) for all drives. Both groups received the same set of HMI screens, including audio-visual feedback, for the transition from MDM to Automated Driving mode) (See Appendix E).

7.3.5.1 CHAT HMI

Participants in the CHAT HMI group received the 'CHAT' HMI as outlined in Table 6-1 and Table 6-2. This consisted of a sequence of audio-visual displays depicting the CHAT SOP behavioural 'checklist' and audio voice messages designed using recent research (Clark et al. 2024) on naturally occurring human-human conversation during a transition of control.

7.3.5.2 Basic HMI

Participants in the Basic HMI group received a baseline HMI design (see Table 6-3). This design was similar to the basic design used in the longitudinal study by Large et al., 2019 and the behavioural training simulator study (see chapter 5). Drivers were provided with an incremental timer bar and a numerical 10 second count down for the final 10 seconds of the PCTP. The MDM symbol gave a visual alert for the driver to place their hands onto the steering wheel.

7.3.6 Procedure

During the study briefing, participants were informed of the vehicle's capabilities and driver fallback responsibilities based on the SAE J3016 (2021) definition of conditional driving automation/level 3 ADS. The drives were framed as a commute to the participant's place of work or study. The aim of this was to shape expectations as to the predictive and familiar nature of the driving route and present a narrative that would encourage an immersive experience and elicit more genuine behaviours from the drivers. As part of this narrative, participants were told to take all of their belongings with them into the car as if they were starting out for the day. Participants were told that they were able to undertake non-driving activities when the ADS was engaged, but that they were required to remain ready to resume manual driving if fallback intervention was required. Participants were explicitly told that the ADS's ODD did not cover the full commute and specifically that it did not cover exiting the motorway. Participants were therefore informed that they would receive a TOR ahead of the junction they needed to take and that they would be given a transition period of 60 seconds before they would need to take back manual control of the vehicle. In addition to the instructions during the verbal briefing, participants were given the opportunity to experience the ADS, the transitions in and out of control and the in-vehicle HMI display sequences during the test drive prior to completing the experimental drives. As part of the briefing and prior to taking part in any of the drives, participants completed a simulator sickness questionnaire (Kennedy et al., 1993) and the Total Trust in Automation Questionnaire (TTAQ) (Gold et al., 2015b).

Participants undertook 4 simulated drives, consisting of a short test drive, followed by 3 experimental drives. Ahead of the test drive, participants were given a briefing on key features of the vehicle and directed to set the cab up so that they were comfortable, e.g. adjusting the seat and mirror. Following the test drive the participant remained in the car while the researcher conducted a brief, verbal well-being check to assess for signs of sim sickness. This consisted of asking the driver to rate how they felt on a scale of 1-10, with 10 meaning they felt fine, and 1 that they felt as though they may vomit (this check was carried out after each of the subsequent drives). The researcher asked if the driver had any further questions and if they were happy to continue. The driver then completed the first of the three experimental drives.

Between each of the experimental drives, drivers left the simulation laboratory and went upstairs to a meeting room for a break of around 10-15 minutes. During this time, they completed the subjective questionnaires, which included the Total Trust in Automation questionnaire (TTAQ) (Gold et al., 2015b); the Situational Awareness Rating technique (SART) scale (Taylor, 2017) and the NASA-TLX workload questionnaire (Hart and Staveland, 1988). The participants

then undertook a cognitive task, presented on a 15-inch Lenovo Thinkpad laptop. The task was designed and presented using Psychopy (Peirce et al., 2019) and consisted of a classic visual search task (Treisman and Gelade, 1980), using two different sets of stimuli. The task took around 5 minutes to complete and was used to provide a cognitive distraction and break from the driving task rather than for any assessment purposes. On completion, participants headed back to the simulation lab for the next experimental drive. Following the final experimental drive, participants completed another simulator sickness questionnaire along with a final set of subjective questionnaires before being debriefed.

7.4 Results

The amount of data recorded during this study is vast, as such the results have been organised into a number of subsections.

Section 7.4.1 covers analysis of the subjective measures reported by the drivers themselves.

Section 7.4.2 analyses drivers actions and responses from the issue of the TOR. This section looks at the response times of drivers to preparing for the takeover, for example, discharging NDRTs, first glance to the roadway etc. It does not cover details on glance behaviour, which is looked at in subsequent sections.

Section 7.4.3 is the largest section and analyses drivers glance behaviour from a number of perspectives (see table for summary descriptions of each section). The reason for taking this approach was to allow analysis of where drivers were focusing their attention throughout the transition period and if this changed depending on particular events or cues from the internal or external environment during the PCTP. It also allowed analysis that linked driver glance behaviour during the PCTP to driver performance following the PCTP to be able to inferences relating to driver SA.

Table 7-5 Summary of glance behaviour analysis by sub-section in 7.4.3

Subsection	Focus of glance behaviour analysis
7.4.3.1	Total PCTP (60 seconds)
7.4.3.2	10 second sequential time periods (Tx) across the PCTP (T1 to T6)
7.4.3.3	Time periods covering the final 20 second of the PCTP (T5 and T6)
7.4.3.4	First 10 seconds following manual driving mode resumption (T7)
7.4.3.5.1	Awareness of lead 2 movements during drive 3 PCTP
7.4.3.5.2	Drive 3 lane change manoeuvre
7.4.3.6	Summary results tables for Glance Behaviour

Section 7.4.4 focuses on analysing driver performance following the PCTP. The metrics used to analyse driver performance within this section were tailored to suit the driving scenarios e.g. lane change behaviour was analysed for the medium and high task demand drives,

whereas speed and lane position were analysed for the low task demand drive.

7.4.1 Subjective Measures

7.4.1.1 Situation Awareness

Driver's subjective view of their SA was determined after each experimental drive using the Situational Awareness Rating Scale (SART) (Taylor, 2017). Driver's ratings are made using 7-point Likert scales, where 1 is labelled 'low', and 7 labelled 'high'. Answers within subscales were combined to evaluate respondents' perception of the attentional demand (complexity, variability, and instability of the situation), attentional supply (division of attention, arousal, concentration and spare mental capacity), and their understanding of the situation (information quality and quantity). The means and standard deviations of both HMI groups are reported in Appendix G and graphically in Figure 7-12 (Low Task Demand), Figure 7-12 (Medium Task Demand), and Figure 7-13 (High Task Demand).

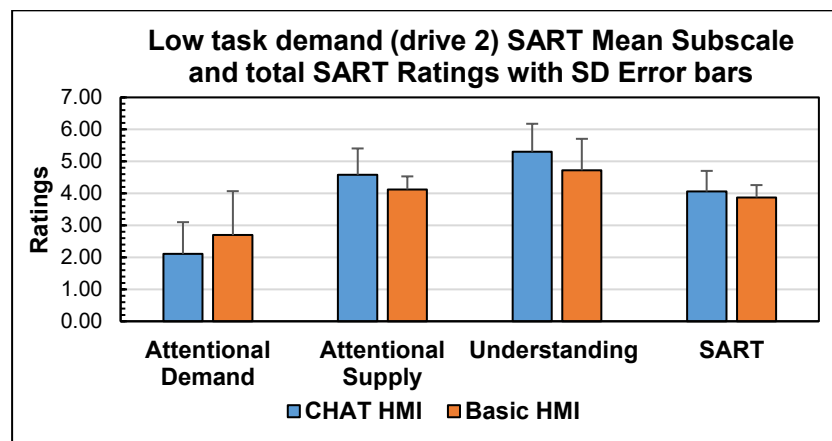


Figure 7-11 Low Task demand (Drive 2) SART subscale and total SART Means and Standard Deviation (SD) error bars for HMI groups (CHAT; Basic).

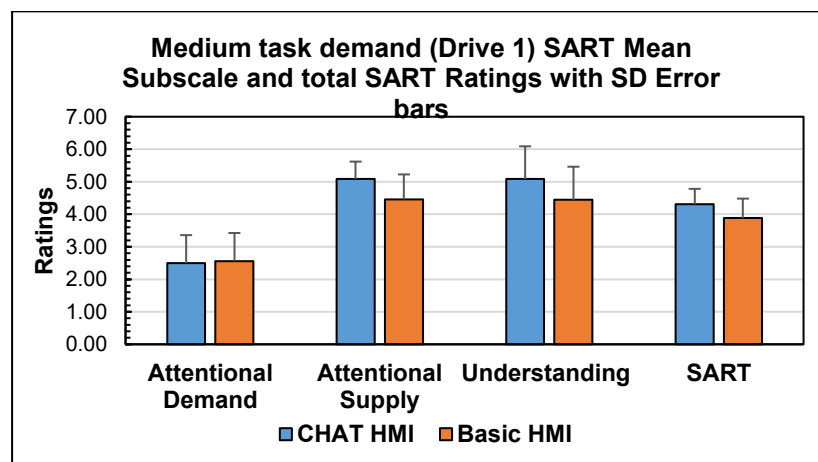


Figure 7-12 Medium Task Demand (Drive 1) SART subscale and total SART Means and Standard Deviation (SD) error bars for HMI groups (CHAT; Basic).

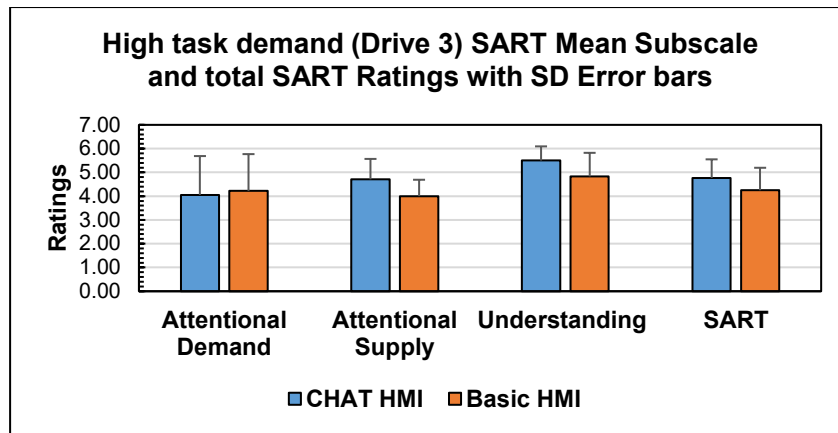


Figure 7-13 High Task Demand (Drive 3) SART subscale and total SART Means and Standard Deviation (SD) error bars for HMI groups (CHAT; Basic).

Attentional Supply

A two-way mixed ANOVA on the subscale of *attentional supply* revealed a significant main effect between HMI groups ($F(1, 22) = 6.420, p = .018, \eta_p^2 = .230$) and task demand ($F(1.872, 41.192) = 5.670, p = .008$). Pairwise comparisons with Bonferroni corrections showed a significantly higher mean ratings for the medium in comparison to the low demand drives ($p = .008$). Mean ratings were higher for the CHAT HMI group compared to the Basic HMI group across the 3 drives and higher for the medium and high task demand drivers compared to the low task demand drive. Mean ratings reported by the Basic HMI group were highest for the medium task demand drive (the first experimental drive conducted and decreased across subsequent drives (the low demand then the high demand drives)).

Attentional Demand

A main effect for the subscale *attentional demand* was also reported for task demand ($F(2, 44) = 27.171, p < .001, \eta_p^2 = .553$). Pairwise comparisons with Bonferroni corrections showed a significant difference between the high demand drive in comparison to the medium demand drive ($p < .001$), and low demand drive ($p < .001$). Mean trends were similar for both HMI groups and highest for the high task demand drive. Interestingly, the Basic group reported higher attentional demand for the low demand drive in comparison to the medium demand drive, whereas ratings for CHAT HMI group matched the task demands of each drive.

Total SART

Finally, a main effect for the *SART total* score was reported for task demand ($F(1.373, 30.200) = 6.646, p = .009, \eta_p^2 = .232$). Pairwise comparisons with Bonferroni corrections showed a significant difference between the low and high demand drives ($p = .003$). Trends show that

mean ratings were higher for the high task demand drive in comparison to the medium and low task demand drives.

Comparisons between the HMI groups

Independent t-tests revealed that following medium task demand drive (Figure 7-12), the CHAT group had significantly higher ratings for *attentional supply* ($t(0.158) = 22, p = .031$) and *total SART* ($t(-1.929) = 22, p = 0.29$) in comparison to the Basic HMI group. Additionally, following the high task demand drive (Figure 7-13), the CHAT HMI group reported significantly higher ratings group for *attentional supply* ($t(0.311) = 22, p = .018$); and *understanding* ($t(-2.000) = 22, p = .29$) compared with the Basic HMI. Effect sizes were calculated using Cohen's d (see Appendix H).

7.4.1.2 Workload

Drivers' subjective views of workload were determined after each experimental drive using the NASA-TLX workload index (Hart and Staveland, 1988). This is a multi-dimensional scale exploring drivers' evaluation of mental, physical, and temporal demand; effort required, and level of frustration (i.e. stress, irritation, insecurity). Ratings were made using 7-point Likert scales, where a rating of 1 represents low workload, and 7, high.

The means and standard deviations (SD) of both HMI groups are reported in Appendix G and graphically in Figure 7-14 (Low demand, Drive 2), Figure 7-15 (Medium demand, Drive 1), and Figure 7-16 (High demand, Drive 3).

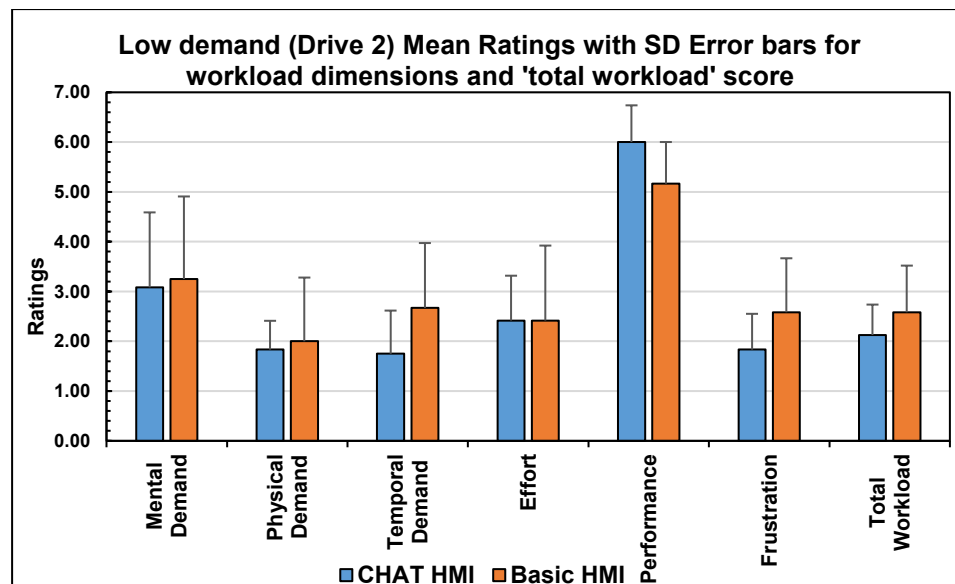


Figure 7-14 Low Task Demand (Drive 2) mean ratings and SD error bars for workload dimensions and cumulative 'total workload' score for each HMI condition.

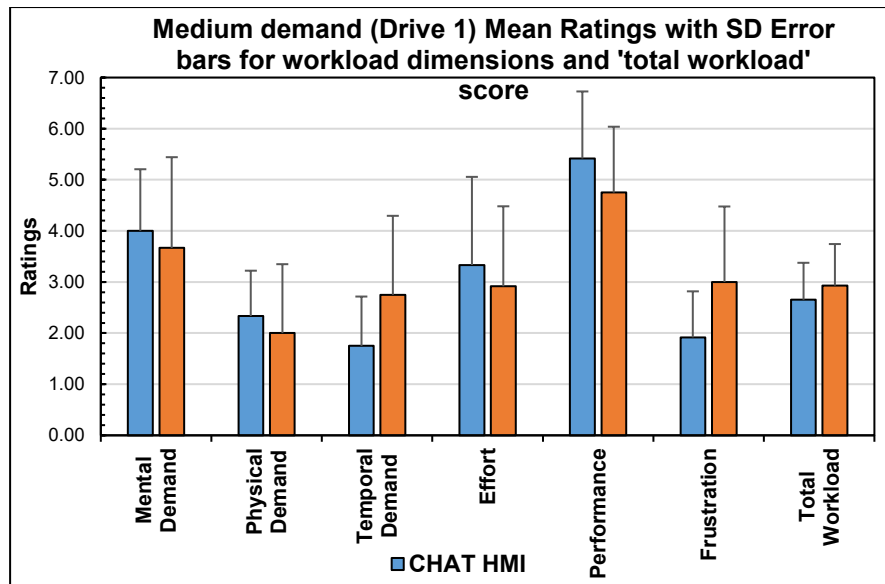


Figure 7-15 Medium Task Demand (Drive 1) mean ratings and SD error bars for workload dimensions and cumulative 'total workload' score for each HMI condition.

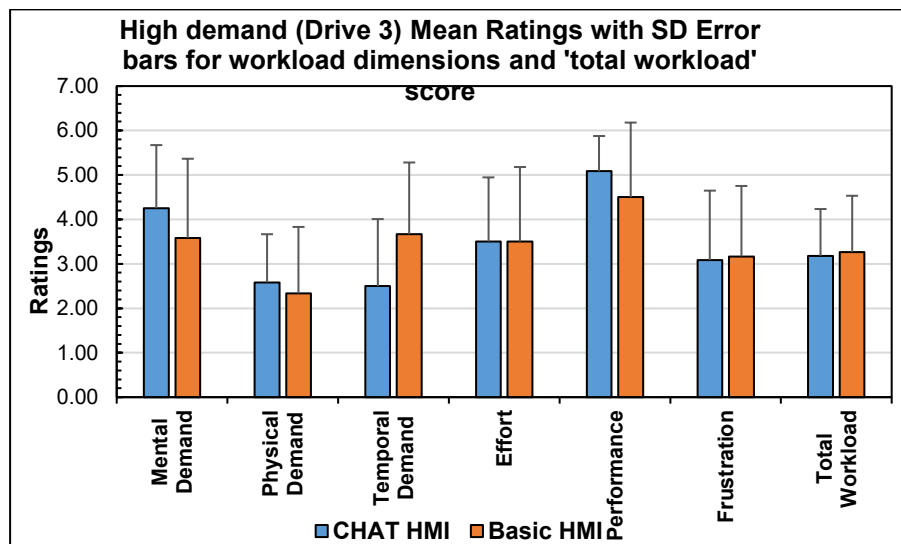


Figure 7-16 High Task Demand (Drive 3) mean ratings and SD error bars for workload dimensions and 'total workload' score for each HMI condition.

Temporal Demand

A two-way mixed ANOVA on the dimension of *temporal demand* revealed a significant main effects for HMI (CHAT; Basic), ($F(1, 22) = 5.748$, $p = .025$, $\eta_p^2 = .207$) and task demand ($F(2, 44) = 7.013$, $p = .002$, $\eta_p^2 = .242$). Pairwise comparisons with Bonferroni corrections for multiple comparisons showed significant differences between the high and low demand drives ($p = .013$) and high and medium demand drives ($p = .026$). Mean trends show higher ratings for this dimension from the Basic HMI group compared to the CHAT HMI group in all three drives.

Mental Demand

A significant main effect was also found for task demand on the dimension of *mental demand* ($F(1.443, 31.742) = 4.412, p = .031, \eta_p^2 = .167$). Pairwise comparisons showed significantly higher levels of mental demand for the medium demand drive in comparison with the low demand drive ($p = .003$). Overall, this suggests that the views of both groups reflected the mental demand required in accordance with the task demands set by the experimental design.

Frustration, Effort and Total Workload

A significant main effect for task demand was found on the dimension of *frustration* ($F(1.847, 44) = 6.447, p = .005, \eta_p^2 = .227$). Pairwise comparisons showed significant differences between the medium and high demand drives ($p = .014$). A significant main effect for task demand was also found for the dimension of *effort* ($F(1.681, 36.976) = 4.739, p = .019, \eta_p^2 = .177$). Pairwise comparisons showed significant differences between low and high demand ($p = .009$). Finally, a significant main effect for task demand was found for the cumulative *total workload* score ($F(1.787, 39.309) = 12.352, p < .001, \eta_p^2 = .360$). Pairwise comparisons showed significant differences between low and medium demand drives ($p = .006$) and low and high demand drives ($p < .001$). Trends relating to how hard drivers felt they had to work to accomplish the level of performance in the task (i.e. effort) matched the level task demand of each drive and this followed the same pattern for frustration and total workload scores.

Comparisons between HMI groups

Independent t-tests revealed that drivers in the Basic HMI group reported significantly higher levels of *temporal demand* compared with the CHAT HMI group in all three drives (low demand ($t(2.030) = 22, p = .027$), medium demand ($t(1.902) = 22, p = .035$), and high demand ($t(2.227) = 22, p = .018$). This suggests that drivers in the Basic HMI group perceived significantly more time pressure in the PCT in comparison with the CHAT HMI group. Additionally, drivers in the CHAT HMI group reported significantly lower levels of *frustration* than the Basic HMI group during the low demand ($t(1.999) = 22, p = .029$) and medium demand ($t(2.169) = 22, p = .021$) drives. Frustration levels increased in the CHAT HMI group following the high task demand drive but were still lower than levels reported by the Basic group (Figure 7-16). This suggests that factors such as stress were higher for the high task demand drive for both HMI groups, compared with the medium and low task demand drives (albeit more so for Basic HMI drivers). Finally, drivers in the CHAT HMI also rated their *performance* (ratings reversed) significantly higher for the low demand drive than drivers in the Basic HMI group ($t(-2.590) = 22, p = .008$). Trend data shows that these drivers viewed their performance more positively than the Basic HMI group

across all three drives, but this difference between the groups was only significant for the low demand drive. Effect sizes were calculated using Cohen's d and are reported in Appendix H.

7.4.1.3 Trust

The total trust in automation questionnaire (TTAQ) (Gold et al., 2015b) was used to explore drivers' views on automated driving and if that changed throughout the study. Descriptive statistics for the Trust subscales can be found in Appendix G. There were no remarkable differences between the two groups in their views on driving automation. A minor point of note is that there was an increasing trend in trust and intent to use until drive 3 and then it reduced slightly, which is unsurprising given the high demand scenario for drive 3. However, in future work, the recalibration of trust during follow up drives might be interesting to compare with different HMI systems.

7.4.1.4 Results Summary – Subjective Measures

Table 7-6 Summary Results for Subjective Measures

Subjective Measure	Summary Results
Situational Awareness (SART)	▪ A main effect was found between HMI groups for <i>Attentional Supply</i>. ▪ The Basic HMI group reported the highest ratings of attentional supply following the 1st experimental drive (medium task demand) then reported decreasing levels across the low demand (2nd drive) and high demand (3rd drive) drives. ▪ The CHAT HMI group reported higher ratings of attentional supply following medium and high task demand drives in comparison to low task demand drive. ▪ The CHAT HMI group reported significantly higher ratings for '<i>Attentional supply</i>', '<i>Understanding</i>' and '<i>Total SART</i>' in comparison to the Basic HMI group following both the medium and high task demand drives.
Workload (NASA-TLX)	▪ A main effect was found between HMI groups and for task demand for <i>Temporal Demand</i>. ▪ The Basic HMI group reported significantly higher ratings for <i>Temporal Demand</i> across all three drives in comparison with the CHAT HMI group. ▪ The Basic HMI group also reported significantly higher ratings for '<i>Frustration</i>' following the low and medium demand drives. ▪ Main effects for task demand were found for '<i>Mental demand</i>', '<i>Frustration</i>', '<i>Effort</i>', and '<i>Total workload</i>'. Generally, these results revealed that the views of both groups reflected the task demands of the drive.
Trust (TTAQ)	▪ No remarkable results to report.

7.4.2 Driver behaviour

A key aim of the study was to evaluate both the necessity and the effectiveness of the CHAT HMI in influencing driver behaviours during the PCTP. The response time data had a non-normal distribution; therefore, non-parametric Mann Whitney U tests were performed to evaluate whether there were differences between the two HMI groups on the following measures:

Time from the TOR onset to:

1. the end of shared attention (i.e. the time the NDRT has been put away and not re-engaged with);
2. the end of the final glance to a NDRT (i.e. the onset time of the first glance following the last coded NDRT glance);
3. placement of hand(s) on steering wheel; and
4. first glance to roadway (a) following the TOR and b) following the end of shared attention).

Results indicated that there were no significant differences between the two HMI groups on any of the above KPIs. However, it was observed that during the drive 1 that several of the drivers in both HMI groups chose not to engage in NDRTs (something that might be expected as drivers develop trust with the automated systems). These drivers tended to remain seated with their eyes on the road and often their hands on the wheel. This has obvious implications for explorations into the behaviour of drivers who have chosen to engage in an NDRT. Therefore, results of no significance are unsurprising. However, one of the proposed benefits of ADS is to allow drivers to engage in NDRTs, thereby maximising time and improving psychological comfort (Carsten and Martens, 2019). As such, investigations into patterns of driver behaviour including whether drivers were choosing to engage with NDRTs during automated driving are important.

A notable observation during this study was the decreasing number of drivers in the Basic HMI group choosing to engage with NDRTs across the drives. By comparison, in the CHAT HMI group, the number of drivers engaging with NDRTs increased. A Chi-squared test revealed that the relationship between the HMI group and the number of drivers engaging with NDRTs was significant $\chi^2 (1, N = 72) = 6.923, p = .009$. Drivers in the Basic HMI group were less likely to engage in the use of an NDRT than drivers in the CHAT HMI group. Looking at the drives individually, there was a statistically significant relationship between the two groups for drive 3 ($\chi^2 (1, N = 24) = 5.042, p = .025$). There was no significant relationship between the HMI group and NDRT use for the first two drives, but the data shows that by the third drive, drivers in the Basic group were significantly less likely to engage with NDRTs during the use of the ADS in comparison to those in the CHAT group. with drivers in the Basic HMI group being significantly less likely to engage

in using a secondary device in the third drive than those in the CHAT group.

Due to this data trend and the fact that the study aimed to evaluate differences in driver behaviour when they engaged in an NDRT during automated driving, an additional Mann Whitney U test was performed between the drivers in each group who engaged with NDRTs.

Descriptive statistics and test results are presented in Table 7-7.

Results are presented in the order the drives were undertaken (rather than the task demand) to show how driver behaviour changed over time (i.e. in relation to the number of drives they had undertaken).

Results indicated that for all three of the experimental drives, drivers in the CHAT HMI had significantly shorter reaction times to end their shared attention following the TOR (i.e. the time it took them to fully discharge their NDRTs) compared to the Basic HMI group. In addition to discharging their secondary activities faster, drivers in the CHAT group also turned their attention away from visually engaging with their NDRTs significantly faster than drivers in the Basic group. Median comparisons showed that drivers in the CHAT HMI group glanced at the roadway following the TOR faster than those in the Basic HMI group for drives 2 and 3. However, differences were not statistically significant, suggesting that the TOR was effective at alerting drivers in both HMI groups to the changing control requirements.

As drivers could respond to the TOR and then return to engaging with their NDRTs an additional measure looked at the time to first roadway glance following the driver's discharge of secondary activities. Drivers in the CHAT HMI group glanced at the roadway sooner than those in the Basic HMI during drivers 1 and 2, and for drive 3, this difference was statistically significant. Finally, for drives 2 and 3, drivers in the CHAT HMI group also placed their hands on the steering wheel statistically significantly sooner following the TOR compared to drivers in the Basic HMI group. It is also worth noting that for the Basic HMI group response times increased for all KPIs across the three drives, whereas, for the CHAT HMI group, response times for KPIs in drive 3 were all lower than in drive 1.

Table 7-7 Descriptive statistics and Mann-Whitney U test results for KPIs (* depicts statistical significance $\alpha = .05$)

KPI	Drive (WL)	CHAT HMI	Basic HMI	Mann-Whitney U	
		Md (IQR)	Md (IQR)	Z	P Value
Time to end of shared attention (discharge of NDRT)	1 (Med)	6.5 (5.87)	7.878 (23.24)	-1.96	.050*
	2 (Low)	2.76 (5.65)	11.625 (13.62)	-2.521	.012*
	3 (High)	3.003 (3.77)	13.623 (19.62)	-2.211	.027*
Time to end of last glance at NDRT	1 (Med)	0.00 (6.19)	7.88 (22.87)	-1.988	.047*
	2 (Low)	0.52 (4.75)	8.825 (13.83)	-2.322	.020*
	3 (High)	4.25 (7.51)	14.37 (20.94)	-2.263	.024*
Time to hands on wheel	1 (Med)	14.88 (8.63)	12.63 (34.13)	-0.159	.874
	2 (Low)	12.30 (9.00)	20.45 (38.56)	-2.064	.041*
	3 (High)	9.26 (11.55)	26.500 (44.63)	-2.111	.035*
Time to first roadway glance after end of shared attention	1 (Med)	7.0 (13.62)	9.88 (22.12)	-1.430	.153
	2 (Low)	2.75 (8.70)	11.7 (12.50)	-1.651	.099
	3 (High)	2.75 (4.51)	12.375 (19.13)	-2.420	.016*
Time to first road glance	1 (Med)	7.0 (8.19)	6.5 (5.39)	-0.371	.758
	2 (Low)	1.25 (6.40)	3.625 (9.13)	-0.908	.395
	3 (High)	3.75 (7.51)	8.00 (14.94)	-1.610	.122

7.4.3 Glance Behaviour

Analysis was carried out to explore any differences in driver glance frequency and duration towards different AOIs grouped under the categories: 'Roadway', 'Internal' and Rear/Peripheral (Table 7-3). Comparisons of visual behaviour were made between the two HMI conditions (CHAT and Basic) and within the different task demand levels (low, medium and high) across the three experimental drives. In addition to the broader 'internal' glance category, driver attention allocation to the HMI display and dashboard was also analysed to provide additional context to driver focus during the PCTP. Analysis was carried on mean glance behaviour for the full sixty second PCTP, which is presented in section 7.4.3.1. This was followed by analysis on the differences between segmented 10-second time periods (T1 – T6) across the PCTP of each drive, to explore how driver behaviour changed across the sixty second transition period and if that changed depending on the drive order and/or task demand. All pairwise comparisons had Bonferroni corrections applied.

Summary tables of key results from glance behaviour can be found in 7.4.3.6.

7.4.3.1 Analysis of the Total PCTP

Mean glance frequency and duration behaviour towards AOI glance categories for the full sixty seconds transition period was analysed. Results are presented by glance category and split by frequency and duration data to analyse differences in visual attention between the two HMI groups and task demand across the full PCTP. Graphs are

presented to show trends in visual behaviour in relation to the different task demands of the experimental drives rather than the order in which the drives were conducted. A summary of results can be found in table x.

7.4.3.1.1 Total PCTP – Roadway

Glance Frequency

Results of a two-way mixed ANOVA evaluating the effect of HMI and task demand on the glance frequency towards the roadway revealed a main effect for task demand ($F(2,44)=17.159$, $p<.001$, $\eta_p^2 = .438$). Pairwise comparisons showed significant differences between the low and medium demand ($p=.008$), low and high demand ($p <0.001$), and medium and high demand ($p = .021$) drives. There was a higher frequency of roadway glances from both HMI groups during the high demand drive (see Figure 7-17).

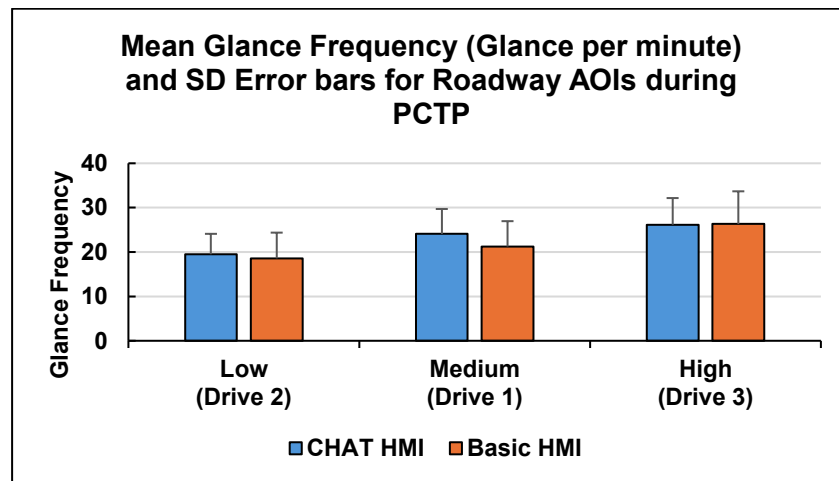


Figure 7-17 Mean glance per minute and standard deviation error bars towards glance category ‘Roadway’ during the PCTP for the low, medium and high demand drives.

Glance Duration

An interaction effect was observed for glance duration to Roadway ($F_{HF}(1.820, 40.049) = 5.256$, $p = .011$, $\eta_p^2 = 0.193$). Simple main effects revealed significant differences for the CHAT HMI group between the low and high demand ($p = .003$), and medium and high demand ($p = .011$) drives, with the high demand drive drawing a significantly higher glance duration towards the roadway than drives 1 and 2. Pairwise comparisons showed no significant differences within the different task demands for the Basic HMI group. Trend data showed that the CHAT HMI group increased the length of glance durations in line with the task demand of each drive, with the greatest level of attention towards the high task demand drive (the final experimental drive). By comparison, the greatest mean glance duration for the Basic HMI was found during

the medium task demand drive, which was the first experimental drive conducted. This group showed similar levels of glance duration for the low and high demand drives.

7.4.3.1.2 Total PCTP – Internal

Glance Frequency

A two-way mixed ANOVA revealed no interaction or main effects for glance frequency to the *Internal* glance category. However, results for the individual AOI 'Dashboard' revealed a main effect between the HMI groups ($F_{HF}(1,22) = 8.725$, $p = .007$, $\eta_p^2 = .284$). Mean glance frequency showed that the Basic HMI group directed higher levels of attention towards the dashboard compared to those in the CHAT HMI group across all drives (Figure 7-18).

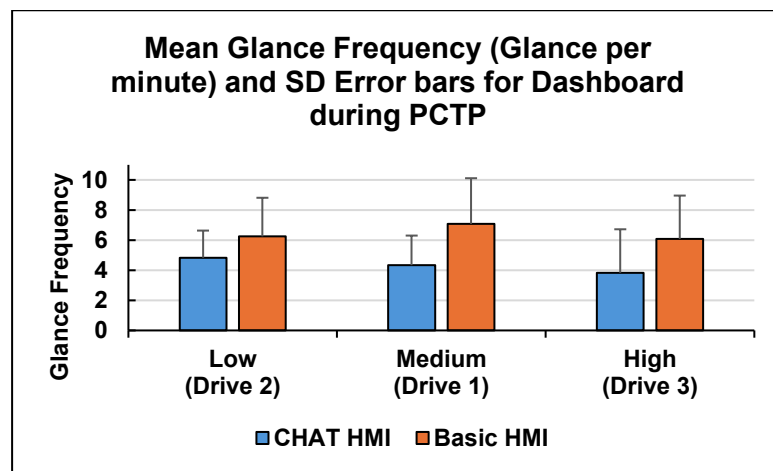


Figure 7-18 Mean glances per minute and standard deviation error bars for Dashboard during the PCTP for the low, medium and high demand drives.

Glance Duration

Results of a two-way mixed ANOVA for glance duration towards the internal glance category, revealed a main effect for task demand ($F_{HF}(2,44) = 5.060$, $p = .011$, $\eta_p^2 = .187$). Pairwise comparisons showed significant differences between the low and high demand drives ($p = .016$). Trends in mean total duration show that the Basic HMI group had higher glance durations towards internal AOIs compared to the CHAT HMI group across all three drives. Trends also show that for the CHAT HMI group, glance duration towards internal AOIs reduced over time across the three experimental drives, with the lowest glance duration towards this area found during the high demand drive (Drive 3) (Figure 7-19).

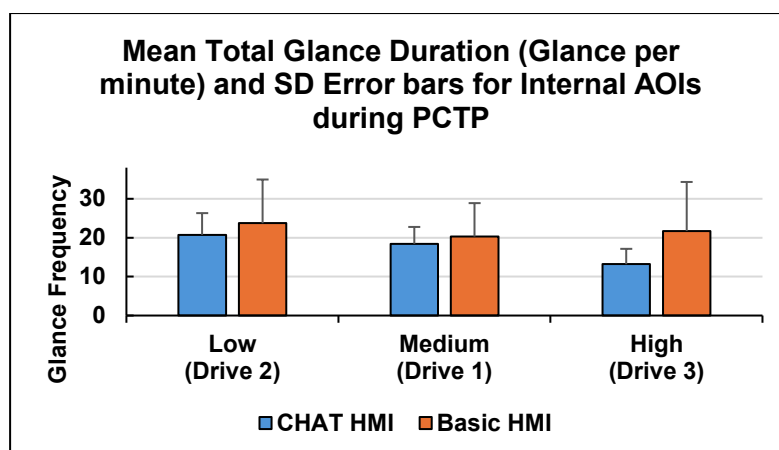


Figure 7-19 Mean glances duration and standard deviation error bars for Internal AOIs during the PCTP for the low, medium and high demand drives.

7.4.3.1.3 Total PCTP - Rear/Peripheral

Glance Frequency

Results from a two-way mixed ANOVA for the glance category Rear/Peripheral revealed a main effect between HMI groups ($F(1,22)=6.801$, $p = .016$, $\eta_p^2 = .236$). Glance frequency towards surrounding road areas was higher for the CHAT HMI group compared to the Basic HMI group for all three experimental drives. No main effect was found for task demand and mean trend analysis shows that glance frequency within each HMI group remained at similar levels for each experimental drive. However, glance frequency to this area did increase for both groups during the high demand drive (Figure 7-20).

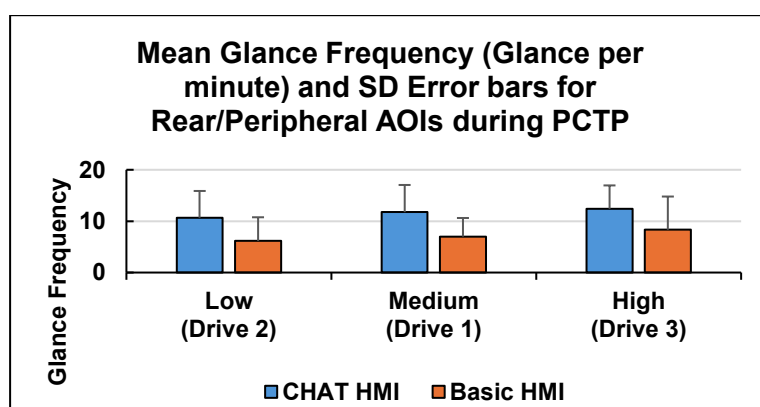


Figure 7-20 Mean glances per minute and standard deviation error bars for Rear/Peripheral AOIs during the PCTP for drives 1, 2 and 3.

Glance Duration

ANOVA results revealed a main effect between HMI groups for glance duration to rear/peripheral AOIs ($F_{HF}(1,22) = 4.772$, $p = .040$, $\eta_p^2 = .178$). Mean trends show a higher glance duration for the CHAT HMI group across all 3 drives (Figure 7-21).

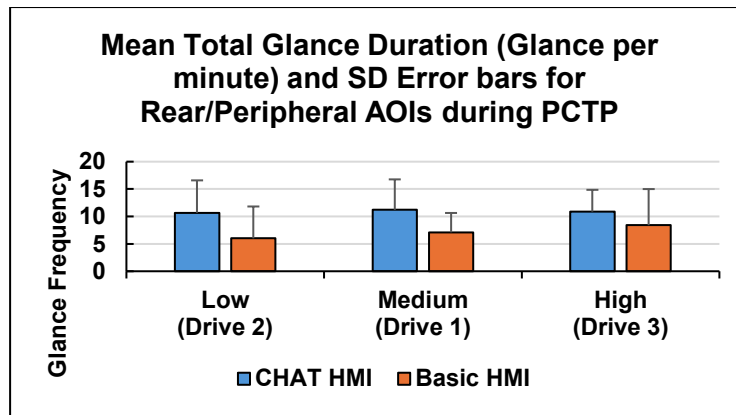


Figure 7-21 Mean duration and standard deviation error bars for Rear/Peripheral AOIs during the PCTP for low, medium and high demand drives.

7.4.3.2 Analysis of PCTP T-periods

The following analysis looks at the visual behaviour of drivers during six segmented ten second time periods across the PCTP for each drive. As with the analysis for total PCTP, results are presented for frequency and duration of glances towards the AOI glance categories of 'roadway', 'internal', and 'rear/peripheral' for each ten second 'T-period' and analyse that in relation to the task demand for the drive as well as events during the drive to present further context for driver behaviour.

7.4.3.2.1 PCTP T-periods – Low Task Demand (Drive 2)

A summary of results from this section can be found in Table 7-19.

7.4.3.2.1.1 Roadway - Low Task Demand

Glance Frequency

Results of the two-way mixed ANOVA for glance frequency towards the roadway for the low task demand drive revealed a main effect between the T-periods ($F_{HF}(5, 95.849) = 3.146, p .011, \eta_p^2 = .125$). Pairwise comparisons showed one significant difference ($p = 0.36$), where glance frequency was lower during the first 10 seconds following the TOR (T1) in comparison to the final 10 seconds of the PCTP (T6).

Glance Duration

A similar main effect was found for glance duration between T-periods ($F_{HF}(5, 109.399) = 7.499, p < .001, \eta_p^2 = .254$). Pairwise comparisons with significant differences are in Table 7-8 below:

Table 7-8 Pairwise comparisons for glance duration towards 'Roadway'

Paired T-periods (T = 10sec)		p value
Pair 1	Pair 2	
T1	T3 (20-30s)	.009
	T4 (30-40s)	<.001
	T5 (40-50s)	.040

Glance durations to the roadway were shorter during the first 10 seconds following the TOR (T1) compared to other T-periods for both HMI groups.

7.4.3.2.1.2 Internal Low - Task Demand

Glance Frequency

A two-way mixed ANOVA for glance frequency to 'Internal' AOIs revealed an interaction effect ($F_{HF}(5, 101.213) = 8.842, p = .041, \eta_p^2 = .287$). Simple main effects with Bonferroni corrections applied, indicated that drivers in the CHAT HMI made a statistically higher mean frequency of glances towards internal AOIs during the first 10 seconds following the TOR ($p = .046$) compared with the Basic HMI group.

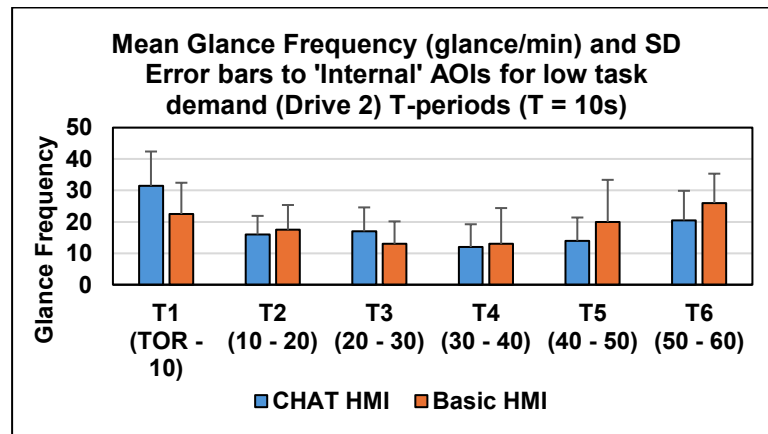


Figure 7-22 Mean glance per minute and standard deviation error bars to 'Internal' AOIs for the low task demand drive (Drive 2) T-periods (T = 10s)

Additionally, statistically significant differences were found between a number of T-periods for each HMI group, summarised in Table 7-9. Mean glance frequency to internal AOIs were significantly higher for the CHAT HMI group during T1 compared to all T-periods except T6. In comparison, internal mean glance frequency for the Basic HMI group was highest during the final 10 seconds of the PCTP (T6), which was significantly higher than the glance frequencies during T3 and T4 (20-40s).

Table 7-9 Statistically significant pairwise comparisons across T-periods (T = 10s) for low demand drive glance duration to 'Internal' ($\alpha = 0.05$).

HMI Group	Pair 1	Pair 2	p value
CHAT	T1	T2	.011
		T3	.026
		T4	<.001
		T5	.006
Basic	T3	T6	.015
	T4	T6	.001

Glance Duration

A two-way mixed ANOVA for glance duration to internal AOIs revealed a main effect for T periods ($F_{HF}(5, 95.804) = 13.938, p < .001, \eta_p^2 = .388$). Pairwise comparisons showed mean glance duration in T1 was significantly greater than all other T-periods. P-values are summarised Table 7-10:

Table 7-10 Pairwise comparisons for glance duration to 'Internal' categories

Paired T-periods (T = 10sec)		p value
Pair 1	Pair 2	
T1	T2 (10-20s)	.002
	T3 (20-30s)	<.001
	T4 (30-40s)	<.001
	T5 (40-50s)	<.001
	T6 (50-60s)	.013

7.4.3.2.1.3 Rear/Peripheral - Low Task Demand

Glance Frequency

Freidman test results for the CHAT HMI group, revealed a statistically significant difference in glance frequency between T-periods, $\chi^2(5) = 18.375, p = .003$. Post hoc tests revealed that mean glance frequency for the CHAT HMI was significantly higher during the T5 (40-50s) than both T1 ($Z = -2.941, p = .003$) and T6 ($Z = -2.625, p = .009$) (see Figure 7-23). Results for the Basic HMI group indicated no significant differences between the T-periods ($\chi^2(5) = 9.071, p = .106$).

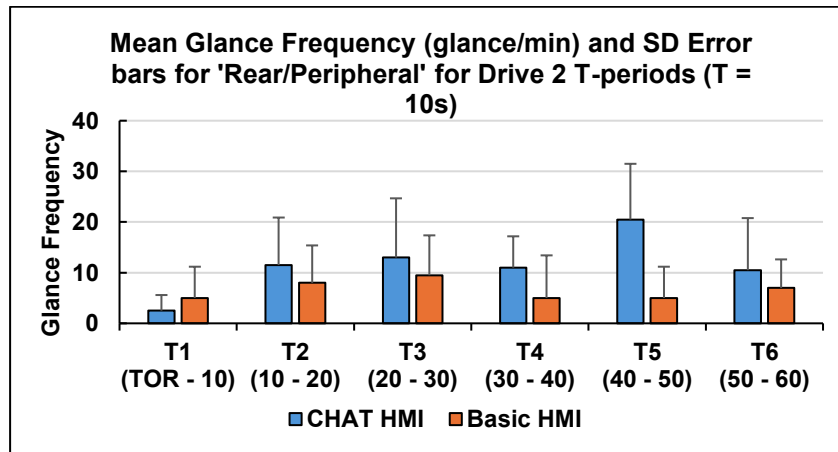


Figure 7-23 Mean glance per minute and standard deviation error bars to 'Rear/Peripheral' AOIs for Low Task Demand drive (Drive 2) T-periods (T = 10s)

Glance Duration

Freidman test results for glance duration revealed a further statistical significance between T-periods for the CHAT HMI group ($\chi^2(5) = 14.939, p = .011$). Post hoc tests revealed a statistically shorter duration

of glances during T1(TOR-10s) in comparison to T5 (40-50s) ($Z = -2.982$, $p = .003$) and T6 (50s – MDM) ($Z = -2.705$, $p = .007$).

Results for the Basic group also showed significant differences between T-periods ($\chi^2(5) = 14.939$, $p = .011$). Post hoc analysis revealed similar results to the CHAT HMI group, with significantly shorter glance durations during T1 compared with T5 ($Z = -2.982$, $p = .003$); and T6, ($Z = -2.705$, $p = .007$).

Glance Frequency and Duration between HMI groups

Mann-Whitney U tests to compare the glance behaviour between the two groups revealed that the CHAT group had significantly higher glance frequency during T4 ($Z = -2.532$, $p = .011$); and T5 ($Z = -3.330$, $p < .001$) and specifically to the rear-view mirror during T5 ($Z = -2.474$, $p = .013$) when compared with drivers in the Basic HMI group. Similarly, the CHAT HMI group had higher total glance duration to rear/peripheral ($Z = -3.492$, $p < .001$) and, in particular, the rear-view mirror ($Z = -2.368$, $p = .018$) during period T5 (40 – 50s) compared with the Basic HMI group. The Basic HMI group showed significantly greater glance duration towards the left-hand mirror during T1 ($Z = 2.134$, $p = .033$), in comparison to the CHAT HMI group. This time period coincided with the ego vehicle passing the entry slip road for junction 29. During this scenario there were no vehicles on the slip road and the ego vehicle remained in lane 1. However, for experimental drive 1 (the medium task demand drive) there were vehicles entering the motorway from this slip road and the ego vehicle moved into lane 2. Glance behaviour suggests that the Basic HMI group's attention was directed towards the slip road during this time but not sustained thereafter.

7.4.3.2.2 PCTP T-periods – Medium Task demand (Drive 1)

A summary of results from this section can be found in Table 7-20.

7.4.3.2.2.1 Roadway - Medium Task demand

Glance Frequency

A two-way mixed ANOVA revealed for the frequency of glances made towards the glance category 'Roadway' revealed no significant interaction or main effects. Trends across the T-periods were similar for both HMI groups, with glance frequency towards the roadway increasing over the course of the PCTP.

Glance Duration

ANOVA test results for total glance duration showed a main effect for T-periods ($F_{HF}(5, 110) = 14.663$, $p < .001$, $\eta_p^2 = .400$). Pairwise comparisons showed significant differences between T1 (the first 10 seconds following the TOR) and all other T-periods. Descriptive

statistics and p-values for significant pairwise comparisons are summarised in Appendix K.

7.4.3.2.2 Internal - Medium Task Demand

Glance Frequency

A two-way mixed ANOVA for glance frequency to 'Internal' AOIs revealed an interaction effect ($F_{HF}(5, 110) = 10.534$, $p = .050$, $\eta_p^2 = .094$). Simple main effects indicated that glance frequency was significantly higher for the Basic HMI group compared with the CHAT HMI group during T4 (30-40s) ($p = .006$) and T5 (40-50s) ($p = .015$). Additionally, statistically significant differences were shown between a number of T-periods for each HMI group, summarised in Table 7-11.

Table 7-11 Statistically significant simple main effects for T-periods (T = 10s) by HMI group for Drive 1 glance frequency to 'Internal' ($\alpha = 0.05$).

HMI Group	Pair 1	Pair 2	p value
CHAT	T1	T3	.018
		T4	.001
		T5	.018
Basic	T1	T2	.016
	T2	T6	.007
	T3	T6	.038

Mean glance frequency for the CHAT HMI group was highest during T1 (TOR-10s), whereas for the Basic HMI group the highest mean glance frequency was during the final 10 seconds of the PCTP (T6). The means and standard deviations of glance frequency for each time period are presented in Figure 7-24.

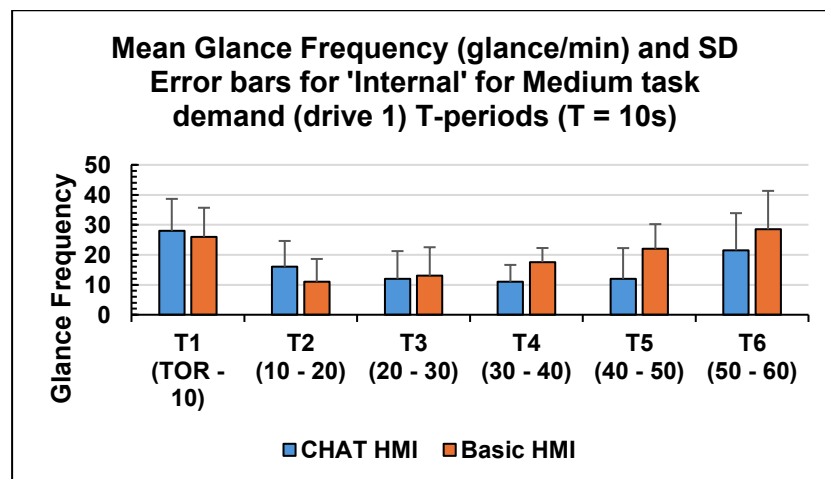


Figure 7-24 Drive 1 mean glance frequency (glance/min) and standard deviation error bars for 'Internal' AOIs during each T period in the PCTP.

The 'internal' category of glances comprised AOIs involved in physically preparing to resume the driving task as well as glances related to managing the transition. Results from a Mann-Whitney U test showed

that the CHAT HMI group made a statistically higher frequency of glances towards the passenger seat during the first 10s following the TOR (T1) ($Z = -2.160$, $p = .031$) and towards their feet or pedals during T2 (10-20s) ($Z = -2.244$, $p = .025$) compared to the Basic HMI group. This suggests that the CHAT HMI group were following the action checklist from the HMI during this time ('put away non-driving activities' (screen 1), 'and get into position for driving' (screen 2)). Both HMI groups showed an increase in glance frequency towards internal AOIs during the final 10 seconds before MDM. However, drivers within the Basic group had a much higher percentage of total glances towards the HMI display during this period compared with those in the CHAT HMT group (Figure 7-25). The medium task demand drive was the first experimental drive conducted by both HMI groups. These results suggests that visual attention allocation towards the HMI display was a lot higher for the Basic HMI group compared to the CHAT HMI group in the final 10 seconds of the PCTP during this drive.

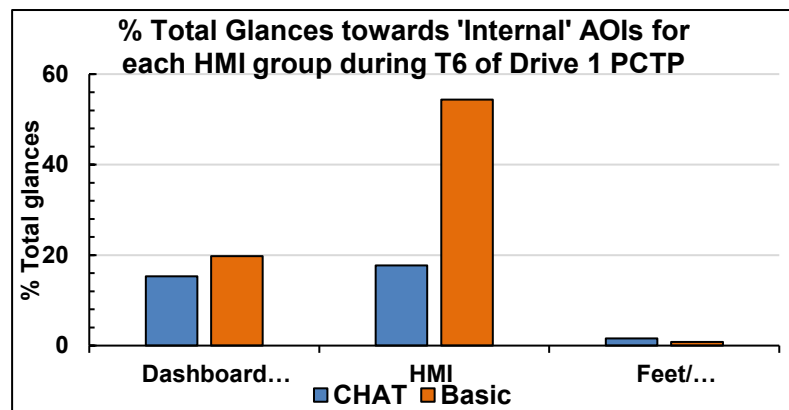


Figure 7-25 % total glances made to internal AOI during T6 (50s – MDM)

Glance Duration

A further two-way mixed ANOVA was carried out to explore the differences in total glance duration towards Internal AOIs across the T-periods. Results revealed a main effect for T-period ($F_{HF}(5, 110) = 22.532$, $p < .001$, $\eta_p^2 = .506$). Pairwise comparisons showed that glance duration to internal AOIs during T1 was statistically higher than each of the other T-periods to the same p-value ($p < .001$).

7.4.3.2.2.3 Medium Task Demand - Rear/Peripheral

Glance Frequency across T-periods

A Freidman test was conducted for each HMI group, to compare differences in mean glance frequency to rear/peripheral AOIs across T-periods (Figure 7-27).

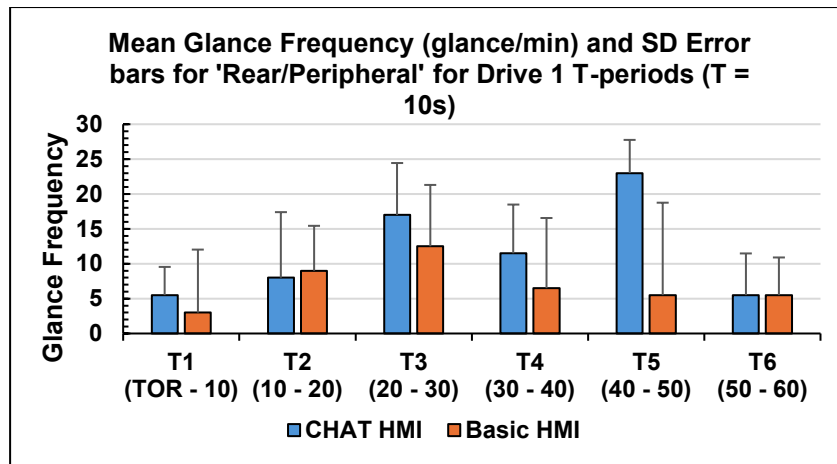


Figure 7-26 Drive 1 mean glance frequency (glance/min) and standard deviation error bars for 'Rear/Peripheral' AOs during each T period in the PCTP.

Results for CHAT HMI group revealed a statistically significant difference in glance frequency between T-periods, $\chi^2(5) = 28.056$, $p < .001$. Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.0083$. The Z and p values for statistically significant differences between T-periods are reported in Table 7-12, with drivers in the CHAT HMI group making a higher frequency of rear and peripheral glances during T5 (40-50s) in comparison to T1 (TOR – 10s), T2 (10 – 20s), and T6 (50s – MDM).

Table 7-12 Statistically significant paired comparisons across T-periods (T = 10s) for Drive 1 total glance duration to 'Rear/Peripheral' ($\alpha = 0.05$).

Paired T-periods (T = 10sec)		Z	p value (< 0.0083)
Pair 1	Pair 2		
T1	T5	-2.952	.003
T2	T5	-2.729	.006
T3	T6	-2.701	.007
T5	T6	-2.829	.005

Results for the Basic HMI group revealed a statistically significant difference for glance frequency between T-periods, $\chi^2(5) = 14.393$, $p = .013$. Post hoc analysis showed a higher mean frequency of glances during T3 (20 – 30s) in comparison with T1(TOR – 10s) ($Z = -2.687$, $p = .007$) and T5 (40-50s) (-2.640 , $p = .008$).

Glance Duration across T-periods

A Freidman test results for the CHAT HMI group revealed a statistically significant different for mean glance duration between T-periods ($\chi^2(5) = 21.496$, $p < .001$). Post hoc Wilcoxon signed-rank tests with Bonferroni correction showed that drivers in the CHAT HMI group had greater total glance duration towards rear and peripheral AOs during T5 (40-50s) in comparison to T1 (TOR – 10s) ($Z = -2.845$ $p = .004$), T2 (10-20s) ($Z = -2.667$ $p = .008$), and T6 (50s – MDM) ($Z = -2.847$, $p = .004$). Results for

the Basic HMI group found no statistically significant differences for glance duration across the T-periods.

Glance Frequency and Duration between HMI groups

Mann-Whitney U tests were performed to compare the differences in mean glance frequency and duration between the two HMI groups at each T-period. Results revealed that the CHAT HMI group made a statistically significant higher frequency ($Z = -3.226$, $p = .001$) and duration ($Z = -3.190$, $p = .001$) of rear/peripheral glances during T5 (40 – 50s) compared with the Basic HMI. More specifically, drivers in the CHAT HMI group made a statistically higher frequency ($Z = -2.244$, $p = .025$) and duration ($Z = -2.121$, $p = .034$) of glances towards the left-hand mirror during T5. The ego vehicle was in lane 2 during the PCTP for drive 1 and, therefore, the driver was required to make a lane change into lane 1 to take the motorway exit. As such, this behaviour suggests that the driver was focusing their attention on the movement of vehicles in relation to required SA for future decision making.

Mean trend data showed that drivers within the CHAT HMI group made more frequent glances towards the rear and peripheral AOIs than those in the Basic HMI group in all T-periods apart from T2 (10-20s) and T6 (50s-MDM), where mean glance frequency was at similar levels (Figure 7-26). Within the scenario, during T2 and T3 (10 – 30s) the ego vehicle approached the entry slip road for junction 29 and pulled into lane 2 to allow a slow-moving vehicle to enter the motorway. Mean glance frequencies towards rear/peripheral areas increased for both HMI groups during this point in the PCTP. However, mean glance frequencies reduced sharply in the Basic group from 30s (T4) onwards. This suggests that the Basic HMI group's attention was directed towards the surrounding road during this time but not sustained thereafter. Mean glance frequencies for both groups were at the same level in the 10 second period leading up to resumption of MDM.

7.4.3.2.3 PCTP T-periods – High Task Demand (Drive 3)

A summary of results from this section can be found in Table 7-21.

7.4.3.2.3.1 High Task Demand - Roadway

Glance Frequency

ANOVA results for Roadway revealed a main effect for T-periods ($F_{HF}(5, 110) = 7.833$, $p < .001$, $\eta_p^2 = .263$). Pairwise comparisons revealed significant differences between T1 (TOR-10s) and the remainder of the T-periods. Trend patterns show lower levels of glance frequency during T1 for both HMI groups (Figure 7-27).

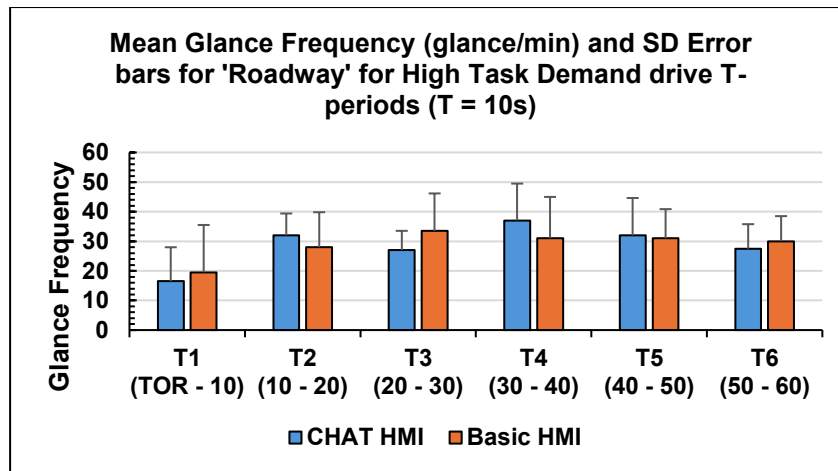


Figure 7-27 Mean glance per minute and standard deviation error bars to 'Roadway' for high task demand drive 3 T-periods (T = 10s)

Glance Duration

A similar main effect was also found for glance duration ($F_{HF}(5, 110) = 9.467, p < .001, \eta_p^2 = .301$) and pairwise comparisons were again found between T1 and the remainder of the T-periods (see Appendix K).

7.4.3.2.3.2 Internal High - Task Demand

Glance Frequency

ANOVA results for glance frequency to Internal AOIs revealed a main effect for HMI ($F(1,22) = 4.787, p = .040$) and for T-periods ($F_{GG}(5, 73.782) = 9.950, p < .001, \eta_p^2 = .311$). Pairwise comparison with Bonferroni correction found significant differences across T-periods as shown in Table 7-13.

Table 7-13 Statistically significant pairwise comparisons across T-periods (T = 10s) for Drive 3 total glance duration to 'Internal' ($\alpha = 0.05$).

Paired T-periods (T = 10sec)		p value
Pair 1	Pair 2	
T1	T2	.027
	T3	.005
	T4	.002
T2	T6	.003
T3	T6	<.001

Figure 7-28 shows that mean glance frequency towards internal AOI was higher for drivers in the CHAT HMI group than the Basic HMI group during T1 (TOR-10), but lower than the Basic HMI group for all other T-periods. The highest mean glance frequency for the Basic HMI group was during the final 10 seconds of the PCTP (T6), whereas for the CHAT HMI group it was during the first 10 seconds following the TOR.

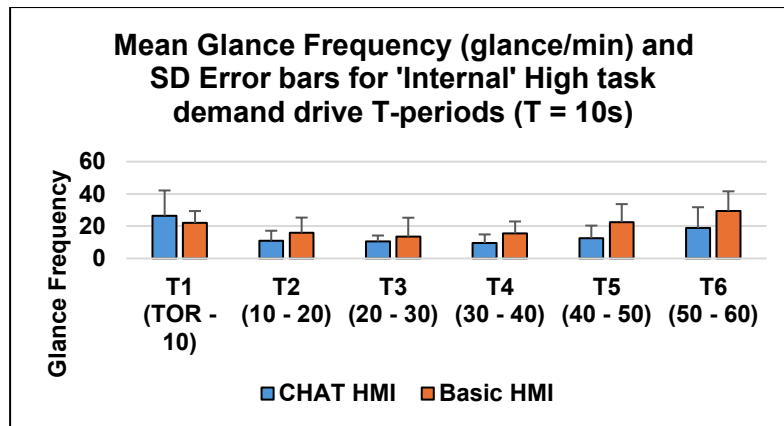


Figure 7-28 Mean glance per minute and standard deviation error bars to 'Internal' AOIs for high task demand drive 3 T-periods (T = 10s)

Glance Duration

A two-way ANOVA for total glance duration towards internal AOIs revealed a main effect for T-period ($F_{HF}(5, 87.273) = 18.122, p < .001, \eta_p^2 = .452$). Pairwise comparison with Bonferroni correction indicated significant differences between T1 and all other T-periods Table 7-14).

Table 7-14 Pairwise comparisons for glance duration to Internal AOIs

Paired T-periods (T = 10sec)		p value
Pair 1	Pair 2	
T1	T2	<.001
	T3	<.001
	T4	<.001
	T5	<.001
	T6	.006

A main effect was also revealed for HMI ($F(1, 22) = 4.899, p = .038, \eta_p^2 = .182$). Mean total glance durations for Internal AOIs were greater for Basic HMI group during T5 and T6 compared to the CHAT HMI (Table 7-15).

Table 7-15 Descriptive Statistics for internal glance duration for T5 and T6

T-period	HMI Group	Mean (s.d.)
T5	Basic	2.85s (1.65)
	CHAT	1.42s (0.89)
T6	Basic	4.24s (2.33)
	CHAT	1.93s (1.54)

7.4.3.2.3.3 Rear/Peripheral - Drive 3

Glance Frequency

A Friedman test for the CHAT HMI group revealed a statistically significant differences in glance frequency between T-periods, ($\chi^2(5) = 18.804, p = .002$). Wilcoxon post hoc analysis indicated that drivers in the CHAT HMI group made a lower frequency of rear and peripheral glances during the T1 than in T2 ($Z = -2.795, p = .005$), T5 ($Z = -2.947, p =$

.003), and T6 ($Z=-2.673$, $p = .008$) (Figure 7-29). No significant differences were found for the Basic HMI group.

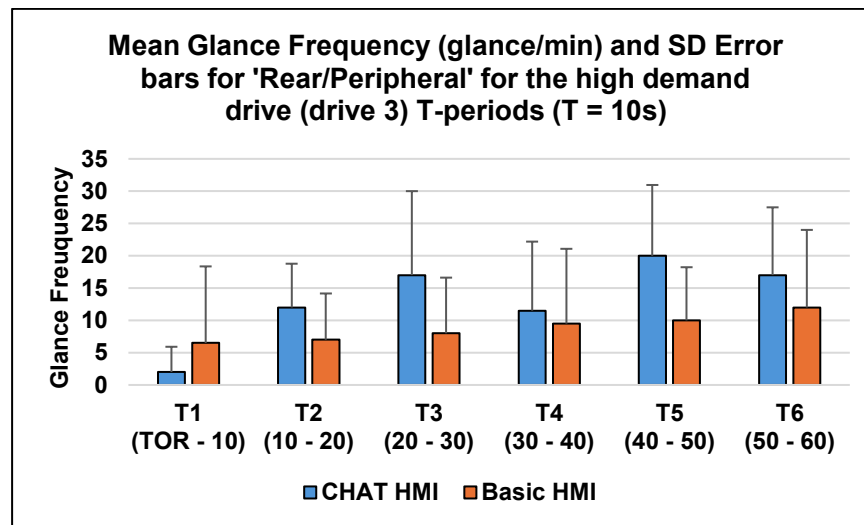


Figure 7-29 Mean glance per minute and standard deviation error bars to 'Rear/Peripheral' AOIs for the high task demand drive T-periods (T = 10s)

Glance Duration

A Friedman test for the CHAT HMI group revealed statistically significant differences in total glance duration between T-periods ($\chi^2(5) = 19.361$, $p = .002$). Post hoc analysis showed similar results to glance frequency with T1 having significantly shorter glance duration than T5 ($Z=-2.934$, $p = .003$) and T6 ($Z=-2.934$, $p = .003$). Again, no significant differences were found for the CHAT HMI group.

Differences between the HMI groups

A Mann-Whitney U test on the differences in glance behaviour towards rear/peripheral AOIs between the two HMI groups revealed that mean glance frequency during T5 (40-50s) was significantly higher for the CHAT HMI group compared with the Basic HMI group ($Z= -2.222$, $p = .026$). More specifically, mean glance frequency ($Z=-2.892$, $p=.004$) and duration (-2.722 , $p = .006$) towards the left mirror during T5 were significantly higher for the CHAT HMI group compared with the Basic HMI group.

Altogether these results suggest that during the high task demand drive, the CHAT HMI group (with the exception of the first 10 seconds following the TOR) had a greater level of attentional allocation towards surrounding road areas across the PCTP compared with the Basic HMI group.

7.4.3.3 Analysis of T5 and T6 % AOI

Time periods T5 and T6 covered the final 20 seconds of the planned control transition period before drivers resumed manual driving control. How long, when and where drivers look gives an indication of the visual

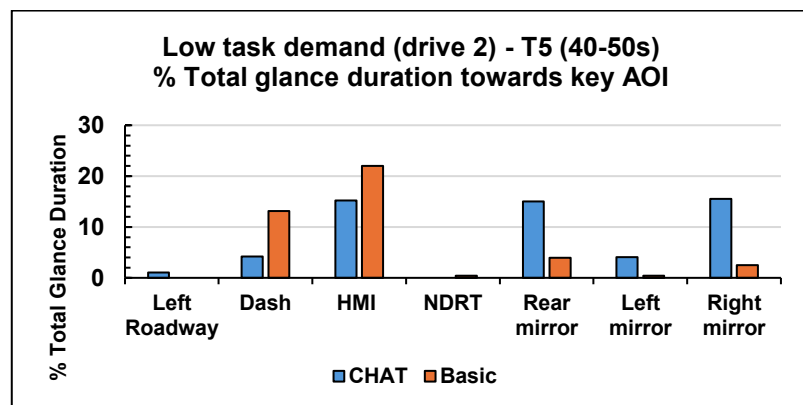
information they are using to build and maintain an accurate representation of the driving environment and form expectations about future events (Seppelt et al., 2017). Therefore, analysis of the percentage total duration of driver glances towards particular AOI during these T-periods would provide a indication of where the drivers were mainly focusing their attention and what information they were concentrating on.

Figure 7-31, **Error! Reference source not found.**, and Figure 7-32, below provide graphs for the low, medium and high task demand drives, respectively, showing the mean % total glance duration towards key AOIs during T5 (40-50s) and T6 (50s – MDM) for each HMI group.

Low Task Demand (Drive 2)

During the low task demand drive, the ego vehicle was in lane 1 for the whole of the PCTP, so the amount required of the driver following the transition period was lower than in the medium and high task demand drives. Drivers from both groups glanced towards their feet/pedals, during the final 10 seconds of the PCTP (T6) and notably, some drivers within the Basic HMI group engaged with their NDRTs during T5 and T6 (Figure 7-30).

a)



b)

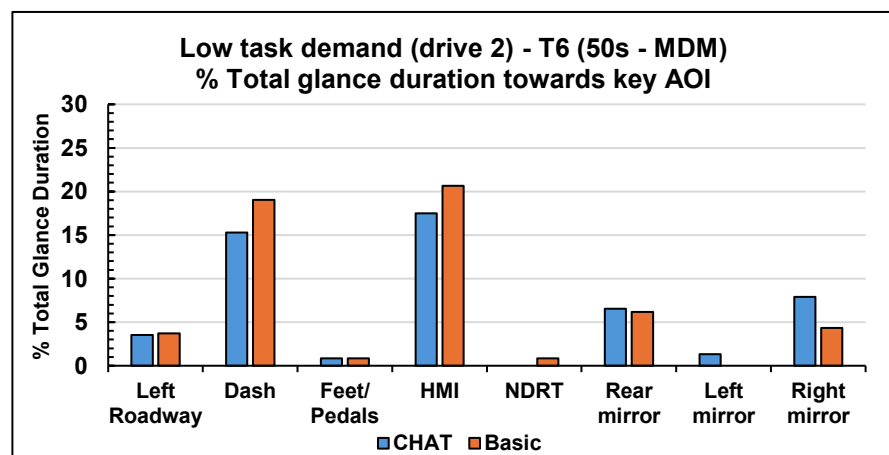


Figure 7-30 Low task demand drive mean % Total Glance Duration towards key AOIs for a) T5 (40 – 50s) and b) T6 (50s – MDM)

There was a noticeable difference in attention allocation (% total glance duration) to internal areas between the two HMI groups during T5, which reduced during T6. For both T-periods, the Basic group gave a higher percentage of total glance duration towards the dashboard and HMI display compared with the CHAT HMI group (Table 7-16):

Table 7-16 % Total Glance duration to Dashboard and HMI during T5 and T6

T-period	AOI	HMI Group	% Total glance duration
T5	Dashboard	Basic	13.1
		CHAT	4.2
	HMI	Basic	22
		CHAT	15.2
T6	Dashboard	Basic	19.3
		CHAT	15.3
	HMI	Basic	20.7
		CHAT	17.5

Comparing across drives, there was a notable increase in the level of attention paid towards the dashboard during T6 by drivers in both HMI groups in comparison to the medium and high demand drives (Table 7-17). This behaviour potentially reflects the lower level of traffic density during this scenario and the lack of requirement to change lanes to exit the motorway.

Table 7-17 % total glance duration for dashboard by HMI group and task demand

HMI Group	Task Demand	% Total glance duration to 'Dashboard'
CHAT	Low	15.3
	Medium	9.5
	High	5.8
Basic	Low	19.3
	Medium	11.9
	High	13.6

Regarding rear/peripheral areas, drivers in both groups gave a similar percentage of total glance duration to the rear mirror during T6 (CHAT , 6.6%; Basic 6.2%). However, drivers in the CHAT group gave a higher level of attention to the right mirror compared with the Basic group during T5 (CHAT 15.5%; Basic 3.9%) and T6 (CHAT 7.9%; Basic, 4.3%). During this part of the scenario, the ego vehicle was in lane 1 and there was a vehicle approaching from the rear in lane 3. Taken together, this behaviour suggests that visual attention was being directed towards the motion of vehicles within the external environment. Notably, this attention was not given by drivers in the Basic HMI and although their percentage of total glance duration increased for the rear mirror, the vehicle in lane 3 was outside of the field of view for this mirror during T6. It could be that mirror glances made by the Basic HMI

group were more arbitrary and by chance rather than directed and purposeful.

Medium Task Demand

During T5 drivers in the CHAT HMI had a higher percentage of total glance duration towards the rear-mirror (CHAT, 9.8%; Basic, 3.9%), left-mirror (CHAT 5.2%; Basic 0.8%) and right-mirror (CHAT, 17.7%; Basic, 2.3%) mirrors compared with the Basic HMI. This difference was particularly pronounced for the right mirror and coincided with a vehicle approaching the ego vehicle in lane 3 at this point in the scenario. Of interest is that drivers in the Basic HMI group did not give the same level of attention towards this area of the external environment when the same cue was present at the same time.

During both T-periods (T5 and T6) the percentages of total glance duration towards the dashboard and HMI display were higher for drivers in the Basic HMI group compared with the CHAT HMI group, especially during T5 (Figure 7-31). Of note is that drivers in the Basic HMI group had a much higher allocation of attention towards the HMI display despite there being a greater degree of visual information on the CHAT HMI display at this point in time. Drivers within both HMI groups made checks to their feet and pedals during T5 and T6.

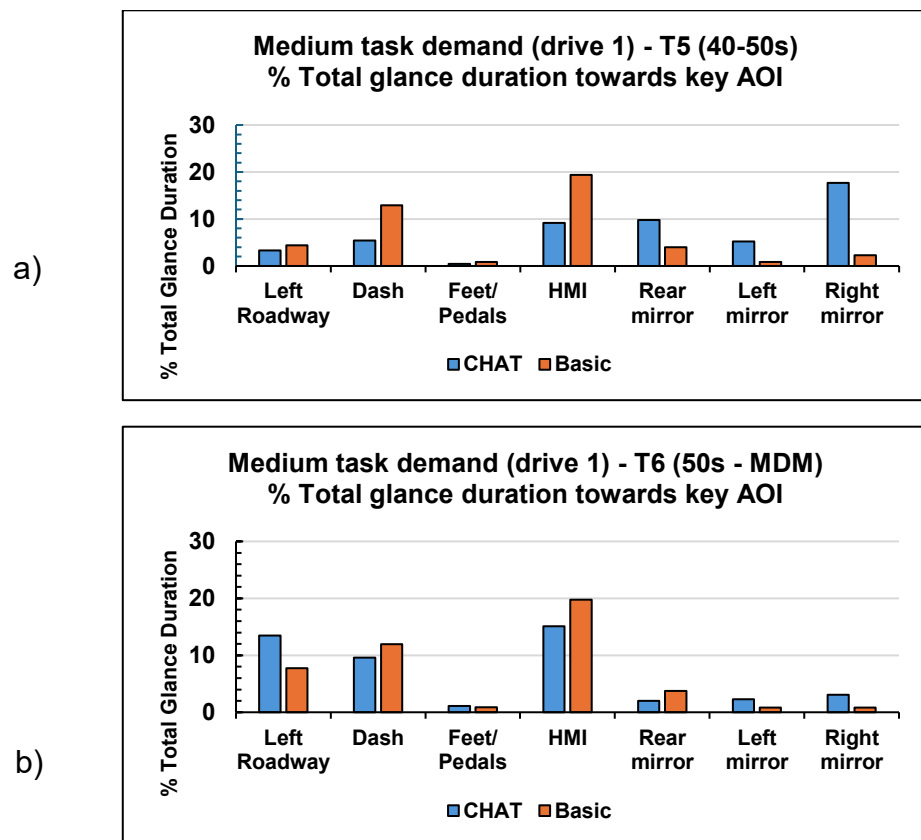


Figure 7-31 Drive 1 mean % Total Glance Duration towards key AOIs for a) T5 (40 – 50s) and b) T6 (50s – MDM)

High Task Demand

During the high task demand drive, the CHAT HMI group predominantly showed a higher percentage of glance duration towards the mirrors than those in the Basic group during T5 and T6 (Figure 7-32). The only exceptions were the rear mirror during T5 (CHAT 11.7%; Basic 13.3%) and the right mirror during T6 (CHAT 2.4%, Basic 2.9%), which were at similar levels between the groups. The higher attention allocation to the surrounding roadway for the CHAT HMI group was particularly apparent during T5, when there were vehicles in both lane 2 (lead 2) (left mirror - CHAT 9.4%; Basic 0.75%) and lane 3 (lead 3) (right mirror – CHAT 9.6%; Basic, 1.7%).

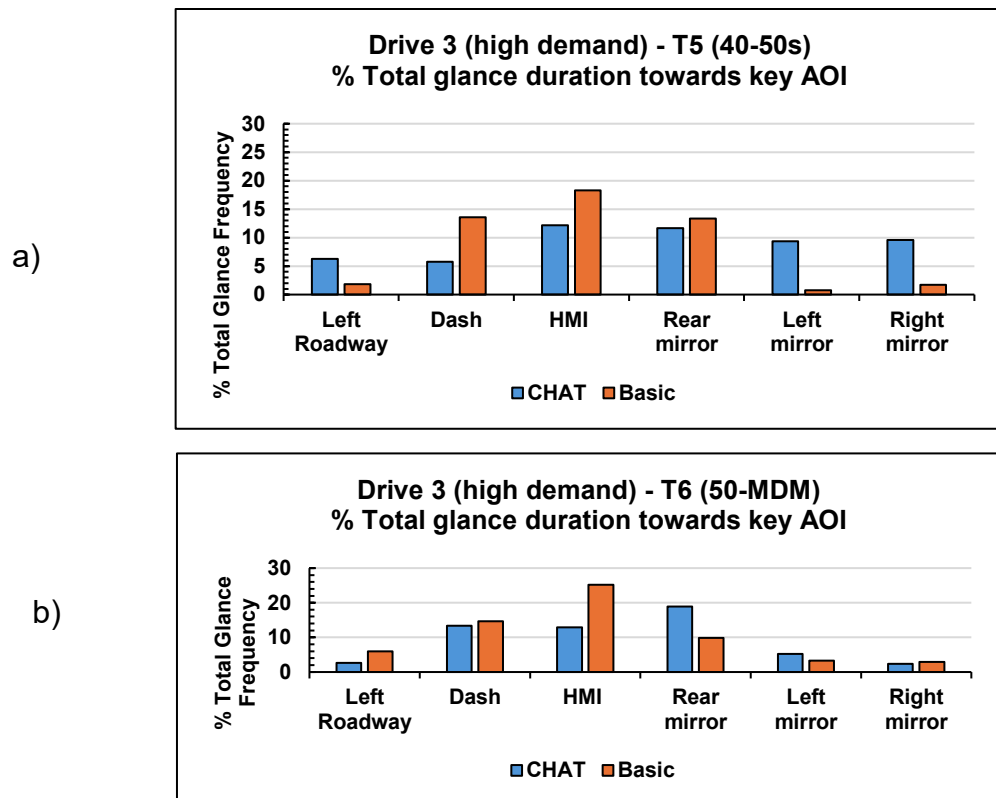


Figure 7-32 Drive 3 mean % Total Glance Duration towards key AOIs for a) T5 (40 – 50s) and b) T6 (50s – MDM)

During T6 the lead 2 vehicle changed lanes from lane 2 to lane 3 and was visible in both the rear and left mirrors. The percentage total of glance duration towards the rear mirror during this time increased for drivers in the CHAT group (from 11.7% in T5 to 18.9% in T6), whereas for drivers in the Basic group it reduced (from 13.3% in T5 to 9.8% in T6), although glances to the left and right mirrors did increase slightly for the Basic group during T6 (from 0.76 to 3.3% and from 1.74% to 2.9% respectively).

There were also differences in attention allocated towards internal AOIs. During T5 drivers in the Basic HMI group had a higher

percentage of glances towards the dashboard and the HMI display compared to drivers in the CHAT HMI group. Interestingly, during T6, drivers in the CHAT group increased their attention towards the dashboard to similar levels of that of the Basic group, albeit still lower. The CHAT HMI group gave similar levels of attention allocation to the HMI display during T5 and T6 and these levels were lower than the % total glance duration during the same periods for the low and medium task demand drives. For the Basic HMI group, attention allocation towards the HMI display increased dramatically from 12.9% in T5 to 25.1% in T6. During the final 10 seconds of the PCTP, this was the AOI with the second largest allocation of attention for these drivers (first was front roadway) in the high demand drive. By comparison, the AOI with the second largest allocation of attention for CHAT drivers during this time was the rear-view mirror (18.9%).

A summary of key results can be found in Table 7-22.

7.4.3.4 T7 – the first 10 seconds post PCTP

Roadway

Results from a two-way mixed ANOVA revealed there were no interaction or main effects for glance frequency or duration towards the Roadway.

Internal

Results of a two-way mixed ANOVA for glance frequency towards Internal AOIs revealed a main effect for task demand ($F_{GG}(2, 31.714) = 9.702, p = .001, \eta_p^2 = .306$) (see Figure 7-33). Pairwise comparisons showed differences to be significant between the low and high demand drives ($p < .001$). For glance duration, results revealed an interaction effect ($F_{HF}(2, 44) = 3.989, p < .001, \eta_p^2 = .306$). Simple main effects found the CHAT HMI group had a significantly shorter mean glance duration towards internal AOIs in the high task demand drive ($p = .011$) compared with the Basic HMI group. The CHAT HMI group also had significantly shorter mean glance duration during the high task demand compared with the low ($p < .001$) and medium ($p = .001$) demand drives.

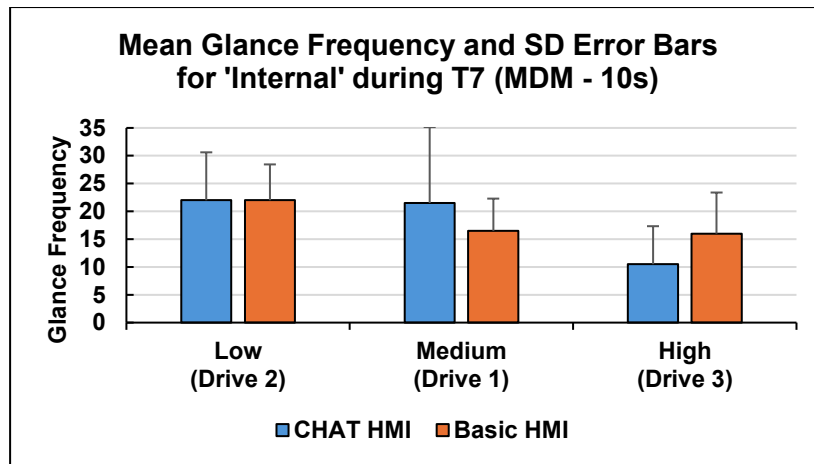


Figure 7-33 Mean glance frequency and SD Error Bars for 'Internal' AOIs during T7 (MDM – 10s)

Mann-Whitney U tests to compare glance behaviour between HMI groups towards individual AOIs within the internal glance category revealed that the Basic HMI group had a significantly higher frequency of glances towards the HMI display during T7 in the low demand drive, ($Z = -2.198$, $p = .028$); medium demand ($Z = -2.198$, $p = .028$); and high demand drives ($Z = -2.460$, $p = .014$) compared with the CHAT HMI group. Similarly, glance duration was significantly higher in the low demand drive, ($Z = -2.460$, $p = .014$); medium demand drive, ($Z = -2.194$, $p = .028$); and high demand drive ($Z = -2.443$, $p = .015$) for the Basic HMI group compared with the CHAT HMI group. In addition, for the high demand drive only, the Basic HMI group also had significantly higher glance duration towards the dashboard ($Z = -2.368$, $p = .018$) compared with the CHAT HMI group (Figure 7-33). This suggests that drivers in the Basic HMI group continued to be distracted by the HMI display after MDM had resumed.

Rear/Peripheral

Freidman tests were performed to evaluate the statistical significance of glance behaviour towards rear/peripheral AOIs for each HMI group. Post hoc analysis with Wilcoxon signed-rank tests was conducted to evaluate this further with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. Results for the CHAT HMI group were significant for both glance frequency ($\chi^2(5) = 19.579$, $p < .001$) and glance duration ($\chi^2(5) = 21.512$, $p < .001$). Post hoc Wilcoxon tests found significantly higher glance frequency and duration during the high task demand drive compared to the low task demand (frequency, $Z = -3.081$, $p = .002$; duration, $Z = -2.747$, $p = .006$) and medium task demand drive (frequency, $Z = -3.084$, $p = .002$; duration, $Z = -2.904$, $p = .004$) (see Figure 7-34). Results for the Basic HMI group were also significant for both glance frequency ($\chi^2(5) = 8.244$, $p = .016$) and duration ($\chi^2(5) = 9.814$, $p = .007$). Post hoc Wilcoxon tests showed a significantly higher level of glance frequency ($Z = -2.386$, $p = .017$) and

duration ($Z = -2.825$, $p = .005$) in the high demand drive compared to the low task demand drive.

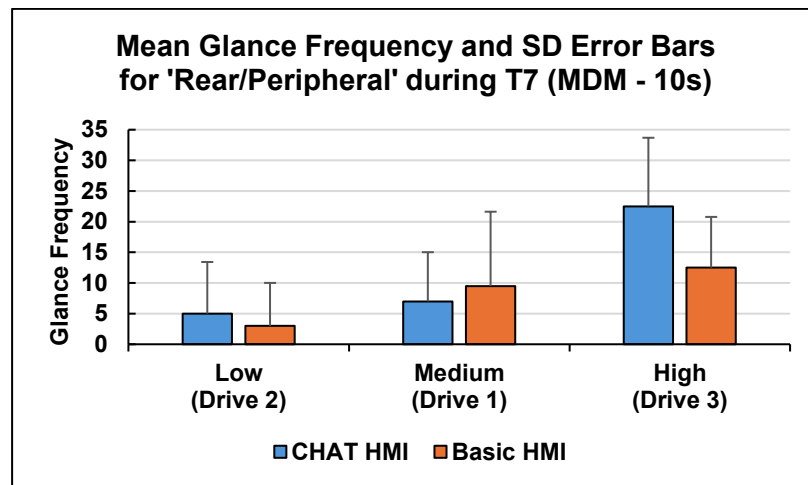


Figure 7-34 Mean Glance Frequency and SD Error Bars for 'Rear/Peripheral' AOIs during T7 (MDM – 10s)

Results of a Mann-Whitney U test to compare the HMI groups revealed that during the medium demand drive, the Basic HMI group showed a significantly higher glance frequency ($Z = -2.138$, $p = .033$) and duration ($Z = -2.135$, $p = .033$), towards the right mirror compared to those in the CHAT HMI group. During the high demand drive, results revealed that drivers within the CHAT HMI group made a significantly higher frequency of glances towards the glance category rear/peripheral ($Z = -2.220$, $p = .026$) and, in particular, the rear mirror $Z = -2.051$, $p = .040$) compared with the Basic HMI group (Figure 7-34).

A summary of key points can be found in Table 7-22.

7.4.3.5 Analysis of the High Task Demand Drive

7.4.3.5.1 Awareness of the Lead 2 vehicle

As part of the study, the high task demand drive (drive 3) was designed to enable an evaluation of drivers' awareness of surrounding vehicles during the PCTP (see Figure 7-35). In particular, key vehicles in relation to the necessary lane change post MDM were of interest. Three vehicles were visible joining the motorway on the entry slip road during the PCTP. The first of these (lead 1) joined the motorway and crossed into the middle lane (lane 2) in front of the ego vehicle. The second (lead 2) and third (lead 3) vehicles joined the motorway behind the ego vehicle during the T4 period of the PCTP (between 40 and 45 seconds after the TOR). Lead 3 crossed into lane 3 to the rear right-hand side of the ego vehicle. Lead 2 moved into lane 2 behind the ego vehicle and during the final 10 seconds of the PCTP moved into lane 1.



Figure 7-35 Schematic showing lead 2 lane change during drive 3

Analysis of drivers' glance behaviour and lead 2 positional data (Figure 7-36) showed that 10 out of 12 drivers in the CHAT HMI (83.3%) glanced to the rear or left mirror when the Lead 2 vehicle was either changing lanes (crossing the lane 1/lane 2 boundary) or positioned in lane 1 prior to MDM engagement. This was compared to 4 out of 12 drivers in the Basic HMI group (33.3%). Additionally, 5 (41.6%) of the drivers in the CHAT group made repeated glances towards the rear or left mirrors during the movement of lead 2 (between crossing the boundary cross and stabilising in lane 1), compared to 1 of the drivers in the Basic HMI group (8.3%). Of note is that there were 2 drivers, one in each HMI group, who made a glance to the rear-view mirror when the onset of the lead 2 lane change had begun (the start of a continuous lateral movement towards the adjacent lane without reversal of lateral position). However, as the lead 2 vehicle had not reached the boundary of lane 1 and these drivers did not return for a second glance, this data was not counted as a positive indication that they had seen lead 2 change lanes and was therefore excluded.

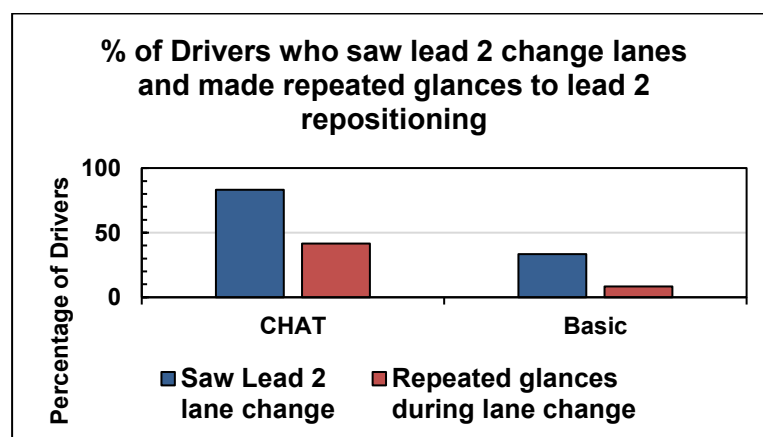


Figure 7-36 % of drivers who glanced towards the mirror during the lead 2 vehicle change lanes and % of drivers who made repeated glances towards the mirror during or following the lead 2 lane change.

A Chi-squared test was employed to evaluate the relationship between the HMI group and the number of mirror glances when lead 2 was changing or had changed lanes. Results showed that the relationship was significant ($\chi^2 (1, N = 24) = 6.171, p = .013$), with drivers in the CHAT HMI group more likely to 'see' the L2 vehicle than drivers in the Basic HMI group. Additionally, a Mann-Whitney U test was employed to compare the differences in the total glance duration towards the rear and left mirrors during the lead 2 lane change and subsequent onward movement in lane 1 (before MDM engaged). Results were not significant for glance count ($Z = -1.914, p = .056$) but show that total glance durations were significantly longer for drivers in the CHAT HMI (mean rank, 15.33) than drivers in the Basic HMI group (mean rank, 9.67) ($Z = -2.038, p = .042$).

7.4.3.5.2 Lane Change

Independent t-tests were used to compare mirror glance behaviour of the drivers in the two HMI groups in the following time periods:

- 10 seconds prior to the lane change onset (T10)
- During the lane change
- Total glances between the start of the 10 second period and the point the left edge of the vehicle met the lane boundary.
- Total no. of glances during the full period.

Results of the t-tests revealed that drivers in the CHAT HMI group had a statistically higher number of glances than those in the Basic HMI group during all measured time periods. Descriptive statistics and test scores are shown in Appendix M and effect sizes in Appendix H. A bar chart of glance count (including the converted count score for T10) is shown in Figure 7-37, for a visual depiction of the differences between the two HMI groups. Together these results suggest that drivers in the CHAT HMI group allocated significantly more attention to the surrounding roadway in preparation for the lane change than those in the Basic HMI group.

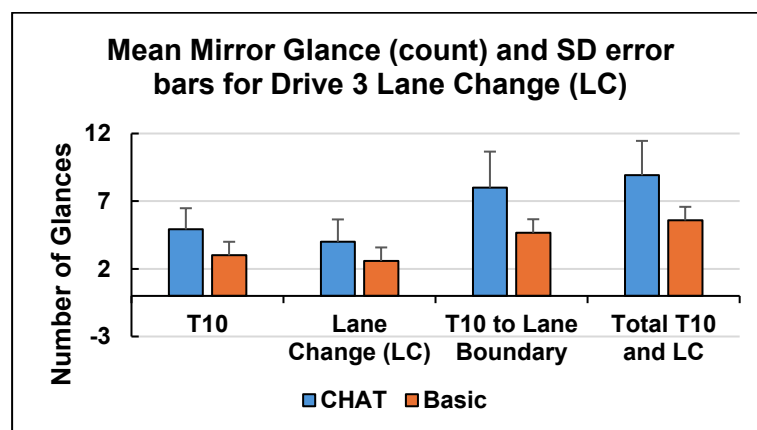


Figure 7-37 Mean count of mirror glances with standard deviation error bars for the drive 3 lane change.

7.4.3.6 Summary results for Glance Behaviour

The following tables provide a summary of results and key points for each analysis section for driver glance behaviour:

Table 7-18 Summary Results for Total PCTP

Glance Category	Summary results for Total PCTP
Roadway	▪ A main effect was found for glance frequency across task demand. ▪ Mean glance frequency matched the level of task demand of the drive for both HMI groups (i.e. the frequency of glances increased with the task demand). ▪ An interaction effect was found for glance duration. ▪ CHAT HMI group showed significant differences in glance duration between the different drives that matched the level of demand for the scenario. ▪ Mean total glance duration was greatest for the high task demand drive for the CHAT HMI group, whereas it was greatest for the medium task demand scenario (the first experimental drive) for the Basic HMI group. ▪ The Basic HMI group showed no significant differences in glance duration between the different drives.
Internal	▪ No interaction or main effects were found for the 'internal' glance category. ▪ However, a main effect was found between HMI groups for the AOI 'Dashboard' (speedometer). Mean glance frequency towards this AOI was higher for drivers in the Basic HMI group compared with the CHAT HMI group for all three drives. ▪ A main effect was found for glance duration across task demand with significant differences between the low and high demand drives. ▪ Mean glance duration was higher for the Basic HMI group compared with the CHAT HMI group for all three drives. ▪ Mean glance frequency reduced over time for the CHAT HMI group with the shortest glance duration during the high task demand drive. It stayed at similar levels for the Basic HMI group for all three drives
Rear/ Peripheral	▪ A main effect was found between HMI groups. Mean glance frequency was higher for the CHAT HMI group compared with the Basic HMI group for all three drives. ▪ No main effect was found for task demand. However, mean glance frequency did increase for both HMI groups during the high task demand drive ▪ A main effect was found between HMI groups. Mean total glance duration was greater for the CHAT HMI group compared with the Basic HMI group for all 3 drives.

Table 7-19 Summary Results for PCTP T-Periods – Low Task Demand

Glance category	Summary Results for PCTP T-Periods: Low Task Demand drive
Roadway	A main effect was found for glance frequency and glance duration for T-periods. For both HMI groups, the mean glance frequency and duration was lowest during the first 10 seconds of the TOR.
Internal	- An interaction effect was found for glance frequency. Mean glance frequency was significantly higher for the CHAT HMI group during T1 in comparison to the Basic HMI group and significantly higher during T1 than all other T-periods across the PCTP. For the Basic HMI group, mean glance frequency was the highest during the final 10 seconds before MDM (T6). - For both HMI groups, mean total glance duration was significantly higher during T1 compared with all other T-periods.
Rear/ Peripheral	- Mean glance frequency for the CHAT HMI group were highest during T5 (40-50s) and significantly higher during this 10-second period than in the first and final 10 seconds of the PCTP. There were no significant differences in mean glance frequency for the Basic HMI group. - Drivers in the CHAT HMI group had significantly higher mean glance frequencies for rear/peripheral and the rear-mirror AOI, during T4 and T5 (30-50s), in comparison to the Basic HMI group. - Mean glance duration for the both HMI group was shortest during T1 (TOR-10s) and significantly shorter during this T-period in comparison to T5 and T6 (40s-MDM). - Drivers in the Basic HMI group had significantly greater total mean glance duration towards the left-hand mirror during T1, which coincided with passing the motorway entry slip road. This suggests that driver attention within the Basic HMI group was directed towards the slip road during this time but this attention was not sustained. In comparison, during T1, driver attention within the CHAT HMI group was on internal AOIs related to physical activities to prepare to drive.

Table 7-20 Summary Results for PCTP T-Periods – Medium Task Demand

Glance category	Summary Results for PCTP T-Periods Medium Task Demand drive
Roadway	- No interaction of main effects found for mean glance frequency. - A main effect was found for total glance duration for T-periods. Mean total glance durations were significantly shorter during T1 (TOR-10s) in comparison to all other T-periods.
Internal	- An interaction effect was revealed for glance frequency. - Mean glance frequency for the CHAT HMI group was highest during T1 and significantly higher during this T-period compared with T3, T4 and T5. - Mean glance frequency was significantly higher for the Basic HMI group in comparison to the CHAT HMI group during T4 and T5 (30-50s).

Glance category	Summary Results for PCTP T-Periods Medium Task Demand drive
Internal (cont.)	- Mean glance frequency for the CHAT HMI group was statistically higher towards the passenger seat during T1 and feet and pedals during T2 compared with the Basic HMI group. This suggests the CHAT HMI group's attention was on physically getting ready for the drive during the first 20 seconds of the PCTP. - A main effect was found for total glance duration for T-periods. Glance durations during T1 were greater in comparison to all other T-periods for both HMI groups.
Rear/ Peripheral	- Mean glance frequency and total glance duration towards the rear/peripheral category, and specifically left-hand mirror, was statistically higher for the CHAT HMI group in comparison to the Basic HMI group during T5 (40-50s). During the PCTP for this scenario, the ego vehicle was in lane 2. This behaviour suggests the CHAT HMI group were attending to the surrounding traffic in relation to future lane change required to take the exit following MDM resumption. - Mean glance frequency for both HMI groups increase during T2 and T3 which coincided with vehicles joining the motorway as the ego vehicle passed the slip road. However, for the Basic HMI group mean glance frequency decreased following this T-period and was significantly lower than T3 during T5. By contrast this was mean glance frequency and total duration were highest for the CHAT HMI group during T5.

Table 7-21 Summary Results for PCTP T-Periods – High Task Demand

Glance category	Summary Results for PCTP T-Periods High Task Demand drive
Roadway	A main effect was found for glance frequency and total glance duration across T-periods with significant differences between T1 (where means were smaller) compared with all other T-periods for both HMI groups.
Internal	- Main effects for glance frequency were found between HMI groups and across T-periods. Glance frequency was highest for the CHAT HMI group during T1 in comparison to all other T-periods. For the Basic HMI group glance frequency was highest during T6 (50s-MDM). Mean glance frequency for the Basic HMI group was higher for all T-periods compared with the CHAT HMI group apart from the first 10 seconds following the TOR. - Main effects for total glance duration were also found between HMI groups and across T-periods. Mean glance durations were greater during T1 compared to all other T-periods and greater for the Basic HMI group during T5 and T6 (40s-MDM) compared to the CHAT HMI group.
Rear/ Peripheral	- Mean glance frequency for the CHAT HMI group was significantly higher during T5 (40-50s) compared with the Basic HMI group. Mean glance frequency and total duration towards the left-hand mirror was also significantly greater during T5 compared with the Basic HMI group. With the exception of T1, mean glance frequency and duration was greater for the CHAT HMI group in comparison to the Basic HMI group. - Mean glance frequency and total glance duration for the CHAT HMI group was significantly higher during T5 and T6 (the final 20s of the PCTP) compared with T1. There were no significant differences across T-periods for the Basic HMI group.

Table 7-22 Key Points for T-Periods T5 (40 – 50s), T6 (50s – MDM) and T7 (MDM – 10s)

T-period	Key points
T5 and T6 (40s – MDM)	▪ Some drivers in the Basic HMI group engaged with their NDRT during T5 and T6. ▪ Drivers in both groups glanced towards their feet and pedals during the final 20 seconds of the PCTP for the low demand and medium demand drives. ▪ The Basic HMI group allocated more visual attention towards the Dash and HMI during T5 and T6 compared to the CHAT HMI group. ▪ Both HMI groups allocated a higher % glance duration towards the dashboard during the low demand drive compared with the medium and high demand drives, potentially reflecting the reduced task requirements (i.e. no lane change) during this scenario, thereby allowing an increased focus on speed control. ▪ Attention allocation towards the HMI display was higher during T5 and T6 for the Basic HMI group in all three drives compared to the CHAT group despite a greater complexity of visual information displayed on the CHAT HMI. ▪ The patterns of attention allocation to rear/peripheral AOIs for the CHAT HMI group in during specific scenario events (e.g. a vehicle passing in lane and increased total % glance duration to the right mirror) suggests that this group were adjusting their attention in relation to cues from the external road environment. A lack of a similar attention allocation pattern from the Basic HMI group suggests that drivers were not attending to the events within the external environment in the same way or to the same level as the CHAT HMI group.
T7 (first 10 seconds post MDM)	▪ No interaction or main effects for Roadway ▪ Main effect found for glance frequency to internal AOIs for task demand. Mean glance frequencies were significantly lower during the high task demand drive compared with low task demand drive for both HMI groups. ▪ The CHAT HMI group had shorter mean glance duration to internal AOI for the high demand drive compared with the Basic HMI group and shorter mean glance duration during the high task demand driver in comparison with the medium and low demand drives. ▪ The Basic HMI group had higher frequency and duration of glances to the HMI display during T7 for all three drives in comparison to the CHAT HMI group. This suggests they continued to be distracted by the HMI display after MDM had resumed. ▪ Significantly greater glance frequency and duration was found for rear/peripheral AOIs for the high demand drive in comparison to the low demand drive for both HMI groups. However, during T7 for the high demand drive, the CHAT HMI group had significantly higher frequency of glances to the rear/peripheral category and, the individual rear-mirror AOI in comparison to the Basic HMI group.

Table 7-23 Key points for Events during high task demand

Events during High Task Demand Drive	Key points for Events during high task demand
Awareness of the Lead 2 vehicle	83.3% of the CHAT HMI group glanced to the rear or left mirror when the lead 2 vehicle was changing or had changed from lane 2 to lane 1 prior to MDM engagement, compared to 33.3% of the Basic HMI group. 41.6% of the CHAT HMI group made more than 1 glance towards the rear or left mirrors during the lane change movement of lead 2, compared with 8.3% of the Basic HMI group. - A Chi-squared test evaluating the relationship between the HMI group and mirror glances during or following the Lead 2 movement into lane 1 revealed that the CHAT HMI group were statistically more likely to allocate attention towards the L2 vehicle during or following the manoeuvre than the Basic HMI group.
Lane change	The CHAT HMI group made a significantly higher number of glances than those in the Basic HMI group in the 10 seconds leading up to and during the lane change manoeuvre.

7.4.4 Driver Capability and Control Response

Due to the variation in the post transition scenarios for each experimental drive, it was not possible to consistently measure drivers' takeover performance across the 3 drives using measures, such as speed and lateral position variability (SD of speed and lane position) over the first ten seconds following MDM engagement, which is typically used as a measure of takeover performance (e.g. Large et al., 2019; Eriksson and Stanton, 2017a). For example, as drives 1 and 3 required a lane change manoeuvre following resumption of MDM, it was not possible to compare performance stability for this time period due to a number of drivers executing this manoeuvre within the first ten seconds post takeover. Additionally, a key focus of this study is on the hypothesis that drivers who had no additional guidance from the CHAT HMI would be more likely to show evidence of a loss of control when dealing with a high task demand driving scenario. Therefore, the predominant focus of this final section of analysis lies in the evaluation of evidence from a combination of video, glance and simulator data for evidence of driver error, loss of control or a near crash between the two HMI groups.

Before discussing the analysis for the high task demand drive, some key points of note will be made in relation to the data on driver performance for the low and medium drives.

7.4.4.1 Low task demand drive speed and lane position

Drive 2 required the driver to stabilise the longitudinal and lateral position of the vehicle and continue to follow a lead vehicle travelling at

the same speed in lane one until the exit slip road for junction 29. Despite this, one driver in the Basic HMI group made an erratic lane change manoeuvre to lane 2 (e.g. max SWA = 15.39 (deg); max lateral acceleration = 0.049 (m/s²); SD speed (kph), 1.51; max speed 85.16 (kph)). This driver then accelerated past the lead vehicle and pulled back into lane 1 at a min THW of 0.57 (s) in front of the vehicle. All other drivers remained in lane 1 and behaviour was unremarkable. A t-test (SD speed) and a Mann-Whitney U test (SDLP) were performed on the mean values of the full 10 seconds. Results showed no significant difference between the two groups for SDLP ($Z = -0.635$, $p = .525$). However, for speed variation, drivers in the Basic HMI group ($M 2.51$, $SD 1.74$) showed significantly greater speed variation ($t(2.270) = 14.045$, $p = .020$), than those in the CHAT HMI group ($M 1.29$, $SD 0.65$). This potentially indicates lower levels of control in the Basic HMI group following the transition, but the results are not considered conclusive evidence.

7.4.4.2 Medium Task Demand lane change

The medium task demand drive required the driver to make a lane change in order to take the motorway exit. A range of performance measures were looked at to evaluate the drivers' performance during the lane change (e.g. Adavikottu and Valaga, 2024), which are summarised in Appendix L. Generally speaking, drivers in the CHAT HMI group took slightly longer to make the lane change, had lower levels and variability in speed and had larger minimum THW and DHW in front of the slow-moving vehicle at the point they crossed the lane boundary. However, results of a Mann-Whitney U test were largely non-significant (with the exception of mean and SD of lateral acceleration, and SD lateral speed). Although these are positive results and potentially hints that the CHAT HMI group had greater control in the lane change manoeuvre, any overarching conclusions are relatively inconclusive.

7.4.4.3 High Task demand lane change and 'loss of control'

To manage the high task demand scenario successfully, the driver of the ego vehicle needed to be aware of the movements of the surrounding vehicles. The response to the lead 1 critical scenario was not the main focus of evaluation for this drive. However, metrics were tested to compare the performance of both groups. There were no significant differences in braking response times or any other driving performance metrics in response to the Lead 1 vehicle braking scenario, suggesting that drivers in both groups responded to this hazardous situation similarly. There was a large level of disparity across the interactions between the task demand of the scenario and the behaviour of individual drivers. This meant that the majority of

performance metrics showed no statistical differences, as despite there being some large differences in the mean data, there was also a large amount of variation, in addition to there being a small sample of drivers ($N = 24$). However, Mann-Whitney U tests did find that the metric 'abrupt onsets of brakes' was statistically higher for the Basic HMI group than for the CHAT HMI group ($Z = -2.340$, $p = .019$) suggesting a higher level of evasive braking response in the Basic HMI group compared with the CHAT group. It should be noted that any abrupt onset of brakes that occurred due to braking in response to the lead vehicle hard brake scenario was not included.

Given the (unsurprising) level of variation between individual drivers in response to the high demand scenario, it was necessary to take an approach to analysis that both considered the contextual nature of the dynamic interactions between the driver and the driving task and took advantage of the vast quantity of data produced by the study. The triangulated BORIS, driver video, and SCANeR data, provided information on driver glance behaviour and physical responses (such as startled facial expressions or body responses) that could be cross referenced with vehicle position and performance data. This 'picture' of individual driver performance was used to place the drivers into two categories using Seppelt et al's (2017) definition of a 'near crash'. Drivers who demonstrated control inputs that constituted an evasive manoeuvre were placed in the category of 'loss of control = yes' and those who did not were placed in the category 'loss of control = no'. There were other behavioural subtleties that could be used to differentiate the performance of the two groups, such as the mirror glance behaviour which was reported on earlier. However, for the purposes of this analysis, a hard line was drawn in relation to evidence (or not) of rapid, evasive control inputs relating to speed, steering or braking, to reduce the subjectivity and potential bias of this evaluation.

Five drivers (41.7%) in the Basic HMI group were placed in the 'Loss of control = yes' category, compared with 0 drivers from the CHAT HMI group. All 5 'Basic' drivers demonstrated evidence of rapid, evasive control inputs during the lane change. Four out of the 5 drivers made rapid steering or braking control inputs (or a combination of both), in response to the late detection of the lead 2 vehicle in lane 1. In response to the position of lead 2, one driver accelerated to a maximum speed of 101.07kph (perceived as 'mph' to the driver) (overtaking all vehicles in lane 1) followed by an evasive braking manoeuvre to pull into lane 1 and take the motorway exit. A Fisher's exact test was used to determine if there was a statistically significant association between the categorisation of loss of control and the HMI group. Results indicated that there was a significant association ($p = .037$) with drivers in the Basic HMI group more likely to be associated with the 'loss of control = yes' category than drivers in the CHAT HMI group. A summary

of the key performance metrics for the drivers in the category 'loss of control = yes' is presented in Appendix I.

There were a number of common behavioural themes between these drivers that are worth highlighting. For example, all 5 drivers glanced at the HMI display after MDM had engaged, including one driver who glanced at the HMI during their aborted lane change. Additionally, 4 out of the 5 drivers had total off-road glances in excess of 2 seconds. On more than one occasion, this appeared to be due to extended and multiple mirror glances, where cues in the road environment (e.g. the lead 2 vehicle) had caught their attention and these glances often spanned either side of the point of MDM engagement. Finally, there was a pattern between the number of mirror checks, over the shoulder checks, and knowledge of the lead 2 vehicle's movements in advance of MDM engagement. Two of the drivers had continued to engage in using their NDRT during the PCTP and had made 1 mirror check between them. Additionally, 4 of the 5 drivers demonstrated poor awareness and glance sampling of the driving environment leading up to the precipitating event (i.e. the combination of the lane change and critical safety incident). For example, 3 of the drivers made 2 or fewer mirror checks leading up to and during the lane change. These mirror checks were made at points where the lead 2 vehicle was not visible in the mirror they were checking, due to its close DHW to the ego vehicle. As these same drivers did not make over the shoulder checks to their blind spot these behaviours led to the requirement for rapid, evasive control inputs to avoid a crash.

An example from each 'Loss of control' category has been illustrated in Figure 7-38 and Figure 7-39 below. In the first example (Figure 7-38), the driver made long successive off-road glances to the dashboard and rear and left mirrors, which resulted in them completely missing the lead 1 hard braking. They spotted the lead 1 part way through their lane change, during a further rear and left mirror check. At this point the lead 1 vehicle was also moving to lane 1 and 0.99s THW to the front of the ego vehicle. This precipitated a rapid and evasive steering control input from the driver and resulted in a prolonged lane exceedance into the hard shoulder lane.

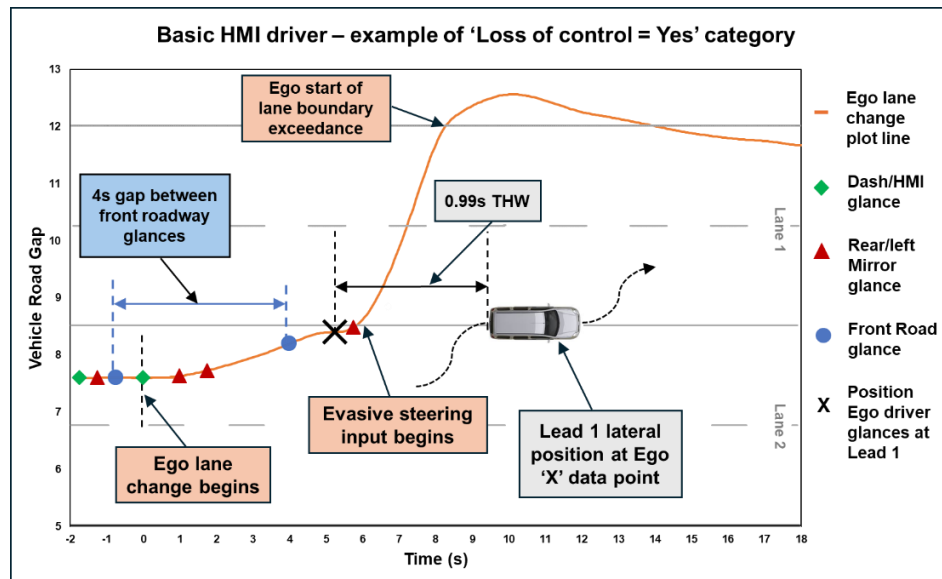


Figure 7-38 Example of a 'Loss of control = Yes' category

In the second example (Figure 7-39), the driver had discharged their laptop early in the PCTP and engaged in checks and assessments of the surrounding environment. They had looked in their mirror at the point lead 2 had changed lanes and returned to glance again when lead 2 was established in lane one. Following the lead 1 hard braking scenario they stabilised their speed and pulled into lane 1 behind lead 2, which they positively confirmed had passed in front of them on the inside lane with two front road glances. It is worth highlighting that three of the drivers within the CHAT HMI group were able to manage the driving task in this way, as they were aware of the positions of the surrounding vehicles including the lead 2 vehicle in lane 1. All three drivers managed the task by stabilising their speed, making additional glances to the mirrors and roadway and then pulling into lane 1 behind the lead 2 vehicle.

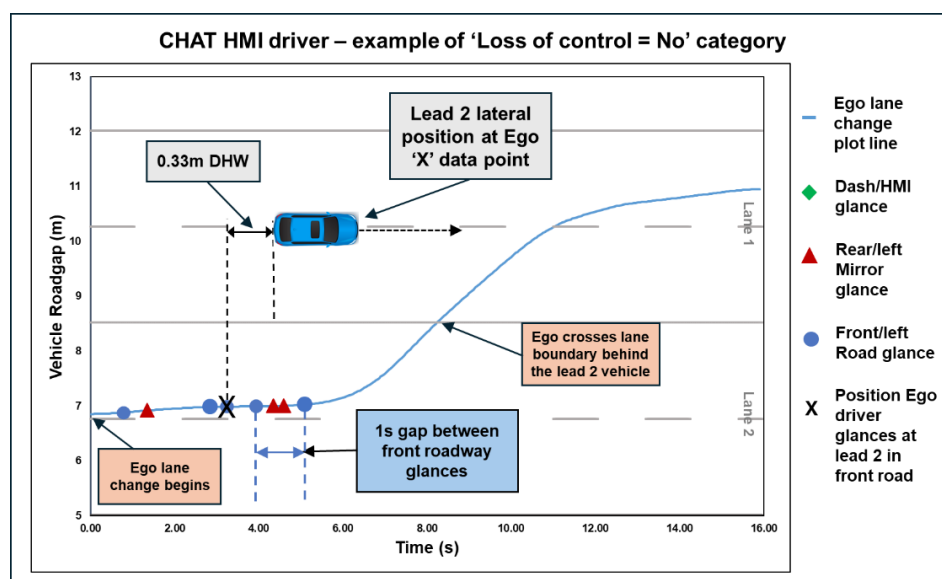


Figure 7-39 Example of a 'Loss of control = No' category

7.5 Discussion

The aim of this driving simulator study was to evaluate the effectiveness of the CHAT HMI design in improving driver behaviour and performance during the PCTP when compared to a Basic (baseline) HMI.

Effectiveness was evaluated across three drives within varying task demands. It was hypothesised that the CHAT HMI would support drivers by reducing shared attention with NDRTs, and increasing SA by improving attentional allocation management strategies and glance scanning patterns across the PCTP. Subsequently, this was expected to facilitate a reduction in driver workload (and increase in psychological comfort) and an increase the likelihood that drivers' available momentary capability would meet or exceed the task demands of the three experimental drives. In particular, during drive 3, it was hypothesised that drivers in the CHAT HMI would be more likely to demonstrate anticipatory control responses towards hazardous incidents. A further aim of the study was to investigate the necessity of additional support or guidance for experienced drivers, or if current skills from manual driving are sufficient to manage the driving task following a takeover. It is proposed by the thesis that PCTs represent a competence gap in the context of driving vehicles equipped with L3 ADS. Additionally, it is argued that drivers, regardless of experience, will not intuitively understand how, when or why to apply manual skills to the automated driving context. Therefore, conversely, drivers provided with the Basic HMI were predicted to be more likely to experience driver errors, loss of control, and/or near crash/crash incidents.

7.5.1 Subjective Situational Awareness, Workload and Trust

Driver ratings of their SA revealed some differences between the two HMI groups. The CHAT HMI group indicated higher levels of 'attentional supply', 'understanding' and the cumulative SART ratings across all three drives, in comparison to the Basic HMI group. Ratings for 'attentional supply' were statistically higher for the CHAT HMI drivers during the medium and high task demand drives (drives 1 and 3 respectively). Additionally, drive 3 garnered significantly higher ratings for 'understanding' from the CHAT HMI group compared with the Basic HMI group. In contrast, the Basic HMI group reported higher ratings for attentional demand, albeit not statistically so. Of note is that the pattern of ratings made by the CHAT HMI group for 'attentional demand' correlated with the experimental drives, with the highest ratings indicated for drive 3 and the lowest for drive 1, whereas ratings from the Basic HMI group indicated a successive increase in attentional demand and decrease in attentional supply across the drives. These results suggest that the CHAT HMI group felt able to give higher levels of attention to the driving task, thereby providing them with a better understanding of the driving task and, in turn, the required attentional

demand (Fuller, 2000; Seppelt, 2017). In contrast, for the Basic HMI group, as their familiarity increased (e.g. with the journey, operations of the vehicle and PCTP), their supply of attentional resources, and, therefore SA calibration, became more difficult (Endsley, 2017a; Mole, 2019). Thus, these drivers' perception of attentional demand for the low demand driving scenario was higher than the medium demand drive they were first exposed to, suggesting a mismatch of attentional resources for the demands of the task (Fuller, 2000).

These observations are supported by the rating responses for workload, which revealed further differences between the two HMI groups. The Basic group reported higher levels of feelings such as insecurity and stress across all three drives compared with the CHAT HMI group. These were statistically higher in drives 1 and 2. Ratings for this 'Frustration' dimension increased for the CHAT HMI group during drive 3, in line with expectations associated with the high demand and safety critical context of this drive. Pertinently, in relation to the evaluation of the HMI designs, the Basic HMI invited statistically higher 'temporal demand' ratings when compared to the CHAT HMI, even during the low demand drive. This is potentially indicative of these drivers allocating insufficient and/or inconsistent attention to important areas of the surrounding roadway during the PCTP. Therefore, important information required to form an accurate expectation of potential future events from surrounding traffic was potentially missed (Seppelt et al., 2017). This reduction in SA would have exposed the drivers to potential spikes in workload and an exaggerated feeling of the pace of task occurrence and a mismatch of task demands (Fuller, 2000).

Exploration in driver trust did not reveal much. Both HMI groups trended more positively for total trust in automation and intent to use scores from the pre-drive through to drive two and then dipped following exposure to the high task demand. These results are unsurprising but may present an area interest for future work relating to the recalibration of trust in subsequent drives and the impact of the HMI used.

7.5.2 Prepare for Takeover

A notable finding from analysis of driver behaviour following the 'prepare for takeover' notification, is the significant relationship between HMI group and the decision to engage with NDRTs. During drive 1, the number of drivers engaging with NDRTs in the Basic HMI group was slightly lower than in the CHAT HMI group. However, by drive three, numbers in the Basic HMI group engaging with NDRTs had decreased, whereas the opposite was true for the CHAT HMI driver population and the relationship between NDRT use and HMI group was statistically significant. Given this trend and the desire to evaluate relative populations (i.e. where drivers in both HMI groups had engaged in NDRTs) analysis was conducted on an isolated population of drivers who use NDRTs during automated driving. Interestingly, findings

showed that the CHAT HMI was particularly effective at directing drivers to discharge NDRTs efficiently and turn their attention towards preparing for the transition ahead. Trends for shared attention went in opposing directions for the CHAT and Basic HMI groups, with the time of shared attention reducing across the drives for the CHAT HMI and increasing for the Basic HMI. This pattern was also shown for time to hands on the steering wheel and the time to the first roadway glance following the end of shared attention. The overarching inferences drawn from this data are, firstly, that the psychological comfort provided by the observability, predictability, directability and timeliness used to design the CHAT HMI (Carsten and Martens, 2019) meant that this group were significantly more likely to engage in using NDRTs (one of the supposed major benefits of higher levels of driving automation); and secondly, these drivers were also far more likely to comply with the request from the ADS to prepare for the takeover.

The TOR alert was the same for both HMI groups. However, the first 30 seconds of the CHAT HMI was dedicated to instructing the drivers to 'put away non-driving activities' and 'get into position for driving'. Therefore, perhaps it is not so surprising that the CHAT HMI shared their attention between the driving and their NDRTs for less time. However, one of the aims of this study was to investigate the necessity for additional driver training or guidance and challenge the assumption that experienced drivers will implicitly understand the required behaviour and actions during a PCT. It could be argued that this is clear evidence in support of this core proponent of the thesis. A further point of note is that both groups were given full agency to behave in a way they felt appropriate or motivated during automated driving. The CHAT HMI issued one audible 'advisement' (Clark et al., 2024) to "put away their NDRTs and get into position for driving". However, this was not accompanied by any non-verbal audio warning cues that may appear irritating or distracting to the driver. Despite this limited 'prodding' from the HMI system and a 30 second lead in time before the "Begin takeover checks" alert, the CHAT HMI group chose to discharge their secondary devices and begin preparing for the takeover at an increasingly shorter timeframe as the drives went on. A reflective point here could be that the conversational manner in which these 'advisements' were delivered to the driver aided the level of collaboration the driver was willing to demonstrate towards the shared driving task during this transition phase.

7.5.3 Check, Assess, Takeover

There were some notable differences in glance behaviour both at the broader PCTP level and when an exploration was carried out on drivers' scanning patterns during sequential 10 second time segments across the PCTP. Generally, drivers' frequency of glances towards the roadway were fairly evenly matched between the two HMI groups.

However, during drive three the front roadway garnered a significantly higher total level of attention from the CHAT HMI group compared to the Basic HMI group, especially during the final 10 seconds before MDM engagement. This increased vigilance towards the front roadway suggests that these drivers were changing their scanning behaviour in accordance with the increased level of traffic movement in their external environment and potential for a precipitating event (Seppelt et al., 2017).

Additionally, analysis of attention allocation to internal and surrounding road areas provides further evidence to suggest the CHAT HMI group were matching their attentional strategies to the demands of the road in a more effective way than the Basic HMI group. For example, in the Basic HMI group, the priority AOI during the PCTP for all drives was speed information, which drew significantly higher glance frequency compared to the CHAT HMI group. Additionally, overall, the total amount of visual attention paid to internal areas, including the HMI display and dashboard, was significantly higher for the Basic HMI group compared with the CHAT HMI group. In contrast, for drivers in the CHAT HMI group, it was the rear and peripheral areas that were prioritised (after visual attention towards the roadway) across all three of the drives. These surrounding roadway areas drawing a significantly higher frequency and duration of glances compared with the Basic HMI group. Incidentally, during drive three, both the total length and rate of glances made by the CHAT HMI group towards off-road, internal areas reduced further, showing a shift in scanning and attentional strategy to increase focus on the front and surrounding road areas. This demonstrates again, not only the potential effectiveness of the CHAT HMI, but also the necessity for an HMI system to support drivers in their management of glance behaviour during this transition period (e.g. Carsten and Martens, 2019; Trosterer et al., 2017, Endsley, 2017a).

Analysis on the differences in glance behaviour between sequential, segmented time periods across the PCTP offered insight into the attentional management strategies employed by drivers, and how these changed over the course of the PCTP, and between each drive. It also offered an opportunity to see whether cues from the HMI or the external environment had an effect on driver behaviour and attention. For example, the number of glances towards internal areas was consistently much higher for the CHAT HMI group in the first 10 seconds following the TOR than the final 10 seconds, following a 'U' shaped pattern across the six time-segments. On closer inspection the makeup of where these glances were focused was different from the first to the last segment, coinciding with the HMI instructions. During the first 10 seconds, drivers were instructed to put away devices and get into position for driving. Driver glance behaviour suggests that they duly complied, focusing on areas such as the passenger seat and their

NDRT as they discharged these devices, followed by their feet/pedals as they repositioned themselves ready for driving, as well as the HMI display as they took in what was being shown and said to them. There was comparatively very little attention towards the dashboard. This weighting towards AOIs within the internal category changed during the final 10 seconds to a greater focus on the dashboard and HMI. It is notable that the HMI display drew significantly more attention from the Basic HMI group than the CHAT HMI group, despite there being more visual information on the CHAT HMI display. During the final time segment of the PCTP the CHAT HMI group tended to increase attention towards the front roadway, and this was interspersed with more frequent glances to internal AOIs and a reduced number of glances towards the surrounding road. This temporal scanning pattern between front roadway and off-road glances to the surrounding areas suggests evidence that these drivers were focused on developing good SA, which could, in turn, increase their capability to anticipate potentially hazardous future events (Seppelt et al., 2017).

From the 30 second 'begin takeover check' point until the final time segment, the attentional strategy of the CHAT HMI drivers was, demonstratively, an increase in focus towards the surrounding road, interleaved predominantly with glances towards the front roadway. Importantly, glances towards the surrounding areas appeared to be deliberate rather than cursory, as the level of attention shown towards the mirrors depended on both the position of the ego vehicle and the salient cues within the external environment. For example, during drive two when the ego vehicle was in lane 1 and there was a vehicle approaching from the rear in lane 3, there was an increase in glances towards the right-hand mirror. In contrast, when the ego vehicle was in lane 2, attentional weight was given either to the left and right mirrors in accordance with vehicles approaching in either lane 1 or lane 3. Notably, this was not evidenced amongst drivers in the basic group, despite these being salient and pertinent objects within close proximity to the ego vehicle shortly before takeover. These findings highlight the importance of focusing on visual attention in transition-of-control research (Pipcorn et al., 2024). They also add weight to the argument that, without explicit guidance, drivers do not view the driving task with a level commensurate with 'shared responsibility' to adequately understand the visual attention required during the PCTP and to develop an accurate picture of current and potential future events (Endsley, 2017a; Mole, 2019).

7.5.4 Driver Capability and Control Response

A key aim of this study was to demonstrate that the available driver capability following the PCT reflected the behavioural choices and attentional strategies adopted by the drivers during the transition period. Higher levels of SA are associated with anticipatory hazard avoidance

(Fuller, 2005; Seppelt et al., 2017), whereas drivers with low SA typically demonstrate more reactive behaviours when dealing with hazards (Endsley, 2018). Thus, the expectation was that patterns of attention allocation during the PCTP would link to patterns in glance behaviour and driving performance post transition, particularly, during the high demand task in drive 3. The consequential effects of not 'checking and assessing' the external environment prior to 'taking over' were prominent in relation to the movements of the lead 2 vehicle in drive three. A significantly higher number of participants in the CHAT HMI group attended to the lead 2 vehicle² compared to the Basic HMI group. Consequently, there was a significantly higher number of the Basic HMI group who did not subsequently take account of the position of the lead 2 vehicle during the lane change and were involved in near crash incidents. Notably, even those participants within the Basic HMI group who did see the lead two vehicle during the PCTP appeared to experience attentional consequences of inconsistent scanning patterns and fewer purposeful (i.e. more cursory) glances towards the surrounding road (as opposed to tracking and assessing traffic movement). For example, drivers whose attention was caught by the movements of the lead 2 vehicle just before MDM engagement subsequently performed elongated off-road glances, neglecting the front roadway (see Figure 7-38 for an example). This resultant 'cognitive tunnelling' effect caused near crash incidents and required evasive steering control input from drivers.

Evidence of the opposite effect was found in the CHAT HMI group, where a number of drivers managed the lane change in a manner that suggested a superior ability to anticipate required responses to events. For example, there were several drivers who, following the emergency brake response to the lead 1 vehicle, stabilised the speed of the ego vehicle and positioned themselves to pull in behind the lead 2 vehicle. Anecdotally, in the debrief following the study, a number of drivers who had not taken this action highlighted that they were concerned that somebody was behind them and so felt compelled to accelerate and pass vehicles in lane 1 before pulling in. In actuality, there were no vehicles positioned to the rear of the ego vehicle at this point so there was opportunity to steady the vehicle and pull into what was the larger gap behind lead 2, if drivers were aware that this was the case. Overall, findings suggest that temporally interspaced and sustained scanning of the road scene and surrounding areas has a positive cumulative effect on anticipatory control responses to hazardous events. In contrast, an attentional strategy absent of these elements may cause the opposite effect as subsequent glances may be made on the basis of incomplete

² Although we cannot definitively prove that these drivers actively perceived the vehicle, their subsequent behaviour such as repeated glances, management of the ego vehicle in relation to lead 2 post takeover and the absence of startle behaviour all provide good evidence to support this inference.

SA (Seppelt et al., 2017). This is further highlighted by evidence that drivers in the Basic HMI group continued to attend to the HMI in the moments following the transition to MDM and increased their off-road glances to an area that bore no relevance to the task at hand.

As predicted, there were very few performance differentials to draw upon post transition for the medium and low demand drives (1 and 2), although there were hints of a reduced level of control amongst the Basic HMI group in comparison to the CHAT HMI group. This demonstrates that the differentials in momentary driver capability between the CHAT and the Basic HMI groups only became truly evident when the task demand was high, and the available capability of the Basic HMI group did not meet the threshold to manage the task difficulty (Fuller, 2000). Both groups of drivers were provided with the same, arguably sufficient, time to raise their SA and attention levels and there were multiple cues within the external environment to draw the drivers' attention if they chose to allocate it. However, without the additional support of the advisements from the CHAT HMI there was a disconnect between the required PCT task and the drivers' understanding of expectation and responsibility of it during that time. It is interesting to note that the CHAT HMI drivers appeared to increase their willingness to collaborate as time went on, perhaps happy that the 'voice' of their 'driving partner' would tell them when it was time to takeover and what they needed them to do (Clark et al. 2024). This was evidenced by the increase in numbers of CHAT HMI drivers who were comfortable to engage in secondary tasks, but also willing (and quick) to discharge them once the TOR was issued. This is in stark contrast to drivers in the Basic HMI group whose comfort in engaging in NDRTs reduced. Moreover, those in the Basic HMI group who did choose to engage in a NDRT, became less willing to disengage from them on receipt on a TOR. Certainly, these promising results provide further support for the use of speech acts and conversational user interfaces for HMI design (Large et al., 2017a; Clark et al., 2024).

7.6 Limitations

Overall, results suggest that the 'CHAT' HMI had a positive influence on glance behaviour and driving performance during and following the PCTP. However, as with any kind of behavioural intervention, success will depend on the willingness and compliance of drivers to follow the instructions. Additionally, a technological solution and SOP offers the potential for consistency of design and a pragmatic way to channel driver behaviour towards desired outcomes. However, its implementation is still reliant on the relevant regulatory and commercial sectors to buy into both facilitating and financing its development and management. For the development of any kind of intervention to have a chance at success, the necessity for the solution will need to be evidenced and validated. It is proffered that the first step in this

challenge has been taken here, but there is still a way to go to validate the findings at scale and offer a persuasive argument that this is a worthy research investment.

As with the previous study, there remains unavoidable, inherent limitations relating to simulator studies, for example, the absence of risk from harm if a collision occurs (Caird and Horrey, 2011). However, much care was taken to address these known challenges and to create as realistic and immersive experience as possible, through the design of the road, movement of the traffic and narrative behind the scenario. A naturalistic approach was taken as far as was practicable, with drivers taking their belongings into the car to provide a range of personal items that they could engage in. However, in order to provide control over the scenario, and to ensure analysis was comparable across drives and participants, the takeover was designed to be passive and provided by the vehicle and not the driver. Future work would need to include a more realistic handover/takeover system, which could give greater authority to the driver to initiate a takeover of the driving task at a time of their choosing. Additionally, as was noted within the discussion section, we can make inferences regarding human perception by cross referencing available data sources, but it is not possible to definitively prove that evidence of attention towards an area where a vehicle is positioned equates to the driver actively perceived the vehicle. Conclusions drawn during the study were validated as far as possible using subsequent behaviour such as repeated glances, vehicular control of the ego vehicle in relation to lead 2 post takeover, and the absence of startle behaviour all provide good evidence to support this inference of perception. However, it is important to acknowledge the limitations as well as mitigate against them. Continued validation of the effectiveness of the CHAT HMI to support driver SA remains a requirement for future work.

Additionally, due to the requirement to ease pressure on recruitment, the three drives were conducted over one day. Even with this concession, issues of available participants for a lengthy study mixed with simulator side effects resulted in a relatively small sample size. Measures were taken in consideration of both of these challenges, for example, taking extended breaks away from the simulation lab between drives and taking a conservative approach in applied corrections for statistical analysis. However, caution should still be applied when drawing absolute conclusions and relating these to on-road behaviour. As with the previous simulator study, the relative validity of the study findings, and the value and potential of the CHAT SOP and HMI concepts designed and evaluated here, are of most importance. It is a fundamental and inevitable necessity to make use of simulated research with safety critical systems that present high risk to

participants, in particular with the exploration of future technology that is not yet available.

7.7 Conclusions

A broad range of evidence has been presented in the results and discussion of this study that demonstrate potential benefits associated with the CHAT HMI design, and support for the concept of the SOP, more widely. Additionally, the behaviour of the Basic HMI group highlights well-known human factors issues associated with a near total disengagement from the driving task during L3 ADS. The potential for driver error and near crash incidents points to the necessity for an intervention to guide driver behaviour during the PCTP. Conversely, the performance of the CHAT HMI group suggests the potential of providing drivers with a simple, timely, consistent and temporally punctuated guidance via a driver display system, as an achievable and pragmatic solution.

Findings raise the importance of an elongated time period prior to the takeover, for interspaced and sustained glance behaviour to be restored (guided by a behavioural intervention). In the context of a PCT, it is posited that asking how much time is needed to take over the driving task is the wrong question in answer to the wrong problem. Instead, it is suggested that, without guidance, drivers typically do not have the required level of available capability following a control transition, to handle sudden high task demand scenarios. Therefore, the question is: how do we adequately and consistently prepare drivers, via education and/or guidance, to encourage the desired (required) driver behaviour during the PCTP to ensure adequate driver capability is available to match the (ongoing) driving task following a takeover. Certainly, we would propose that the concept of CHAT is a step in the right direction in response to this challenge.

8. Discussion and Future Directions

8.1 Introduction

This final chapter provides an overall discussion of the research presented in the thesis. It provides an evaluation of the extent to which the research has answered the research questions and responded to the overall aims, presented in chapter one and summarised in section 10.1.1 below. Within the evaluation, a discussion of the limitations to the research and methodology used will also be addressed. Lastly, this chapter identifies how this thesis contributes to the current topic, by considering the significance and implications of the findings from a theoretical, methodological and practical perspective and future work to build on the results.

8.1.1 Research Aims and Questions

The PhD used a mixed methods approach to examine the necessity for and potential solutions to driver training for experienced drivers using L3 automated driving systems (ADS). There were three core aims, to:

1. identify human factors issues relating to driver behaviour in vehicles with conditional driving automation.
2. investigate driver interactions and behavioural adaptations in the context of L3 'Conditional Driving Automation' and planned control transitions.
3. design and evaluate the effectiveness of training and technological solutions to support desired driver behaviour during planned control transitions in vehicles with conditional driving automation.

These aims were linked to four key research questions, which were disseminated across three phases of work: Phase one ('Background and context') and phase two ('Analysing driver behaviour to guide design') were associated with responding to research questions one and two:

- RQ1. What can we learn from how drivers interact with L3 automated driving systems (ADS) to inform driver training and behavioural requirements
- RQ2. What behaviours and actions are necessary and sufficient to prepare drivers to resume the dynamic driving task (DDT) from automated driving systems during dynamic operations?

The objectives of these early phases of work were to build a foundation of knowledge and data to inform the development of potential design solutions as well as start to build evidence to justify the necessity for behavioural intervention.

Phase three (Design solutions and evaluation) was associated with responding to research questions three and four:

- RQ3. How can driver training support the uptake and maintenance of desired driver behaviour and provide a countermeasure to undesirable behavioural adaptations/operational deviations?
- RQ4. How can technological solutions (e.g. in-vehicle HMI display systems) support an 'on the loop' state in drivers during automated driving to support the calibration of SA and attention for smooth transitions of control?

The aim of phase three was to design, develop and evaluate behavioural intervention solutions that applied the CHAT standard operating procedure to training or technological solutions that could support drivers in undertaking planned control transitions (PCTs). Additionally, part way through this phase, the aims and objectives from phases one and two were expanded upon through an extended review of the literature, and a thematic analysis of driver interviews from the driver training simulator study. This work aimed to reflect on and evaluate the research conducted up to this point, both from a theoretical and methodological perspective, in order to build on the foundation of knowledge that could be applied to the final design solution and evaluation.

8.2 Planned Control Transitions. New task: sufficient competence or necessary training?

Before discussing the effectiveness of the proposed design solutions developed as part of this thesis, it is important to reiterate the justification for the research and the identified research gap. This includes the approach taken to validate the proposed necessity for additional training or guidance for experienced drivers instead of the passive approaches currently offered (i.e. user manuals or dealership instruction). It is worth noting that the importance of research within this area has been further emphasised by a recent report by the Department for Transport (2025). The report emphasises the significance of the challenge faced by drivers in regaining situational awareness (SA) following a period of using automated driving systems, as well as the criticality of regaining SA prior to resuming manual driving to enable a safe takeover.

Future vehicles are expected to be capable of delivering intermediate levels of automation that will allow drivers to relinquish control of the driving task to their vehicle's ADS under certain, predefined conditions. However, drivers of these vehicles will maintain fallback responsibility, meaning that if and when intervention is requested by the system, the driver must be ready to take back control of the driving task. These 'control transitions' have drawn lots of attention within automotive automation research (Eriksson and Stanton, 2017a). However, significant focus has been on 'failure induced' control transitions

(Eriksson and Stanton, 2017b), which typically involve a critical incident as the rationale behind the takeover request (TOR). Takeover response time as well as the ability to handle the critical event that triggered the request are typical metrics by which to measure success (e.g. see Gold et al., 2018 and Eriksson and Stanton, 2017a).

The focus on takeovers due to critical incidents and/or system failures is a logical place for research to concentrate in relation to level 2 ADS, where the predominant paradigm sits with lower system capability and reliability and a driver requirement to continually monitor the system. However, the thesis argues that at L3, system capability and reliability should be such that the predominant rationale for a TOR is due to an approaching boundary of the system's operational design domain. The additional expectation (and indeed assumption made by the thesis) is that most ODD (and enforced, regulatory) boundaries will mean that ADS capability will be unlikely to cover most (if not all) end-to-end journeys. Therefore, 'planned control transitions' (PCTs), will represent a new sub-task for the human driver in the context of vehicles equipped with L3 ADS.

It is the supposition of this research that the fundamental change to the ideology, and associated expectations and responsibilities of the new driving task, specifically, the PCT subtask, means that there is a requirement to clearly impart an accurate mental model of this new joint activity to all drivers (Beggiato and Krems, 2013; Beggiato et al., 2015; Casner and Hutchins, 2019). In particular, the skills, rules and knowledge associated with the new interrelationship and required interactions between the system and the human driver to develop appropriate driver competence (Fuller, 2000). The thesis proposes that the lack of attention towards PCTs represents a research gap worthy of note. Furthermore, the research findings presented within this body of work advocate for the inclusion and importance of this type of control transition and the necessity for drivers to receive effective training and guidance in the new shared driving task. This is especially important for what is a widely used safety critical system, whose operators have a diverse and disparate range of skills and experience. The author suggests that the importance of training or instructional guidance is equally important for experienced drivers as it is for novices, potentially more-so, as this population will be unlikely to undertake further formal training ahead of having access to these vehicles on public roads, and they will likely be the first to use vehicles equipped with these technologies. Appropriate interventions for this group of drivers, therefore, represents a priority for research development within this area.

Despite the skills and experience this category of drivers will have amassed, the context of PCTs in vehicles equipped with L3 ADS requires a fundamentally different understanding of the driving task and

application of these skills. In particular, driver's visual attention allocation for the purposes of rebuilding SA following periods of near complete disengagement from the driving task (Endsley, 2017a; Mole et al., 2019; Carsten and Martens, 2019). As highlighted by Pipcorn et al. (2024), despite the importance of visual attention for vehicle safety, research into control transitions has continued to have a greater focus on take-over times. However, evidence from the studies presented here, as well as research on public roads (Pipcorn et al., 2024), highlights that drivers may takeover control before they are aware of the driving environment. Therefore, it cannot be assumed that drivers are ready to respond to events shortly after the TOR.

Behavioural adaptations of operators of these new joint cognitive systems can introduce a number of issues relating to aspects of performance and safety, as observed when automation was introduced to the aviation domain (Parasuraman et al., 2000; Bainbridge, 1982). Findings from the observation study carried out in chapter three highlighted a number of key behavioural adaptations and operational deviations that provide support for empirical evidence from previous research. Namely, that extended periods of near complete human disengagement from vehicle operation poses a significant challenge to driver attentional allocation and subsequent SA (Carsten and Martens, 2019; Endsley, 2017a; Mole et al., 2019). In particular, as well-learned, and often automatically processed, performance patterns, such as attention allocation through well-rehearsed glance behavioural sequences, are disrupted (Endsley and Kiris, 1995). For example, analysis of driver's visual behaviour across the multiple PCTs found that drivers appeared to develop glance strategies to improve the management of vehicular control at the point of manual driving mode (MDM) engagement. However, they demonstrated little to no attention allocation to the surrounding roadway during the sixty seconds leading up to the takeover, suggesting either an automation bias (or complacency), i.e. an over-reliance on the ADS, or certainly, a lack of understanding of the expectations and responsibility in relation to their role in the shared driving task during this period of transition to manual driving. It also indicates a lack of understanding of the importance and necessity of carrying out behaviours and actions to rebuild levels of SA and attention ahead of resuming control of the driving task (and potential consequences of not) regardless of the provision of sufficient time to do so (Carsten and Martens, 2019; Mole et al., 2019; Pipcorn et al., 2024).

This pattern of inattention was again seen with drivers who did not receive any behavioural intervention in the driving simulator studies described in chapters five and seven. This provides further validation for the necessity for instructional guidance or training to be provided to drivers to shape desired behaviour during these transition periods. This

was particularly pertinent with the findings presented in chapter seven, as a number of steps had been taken to increase the face and ecological validity of this study. The aim of the improvements to the experimental design were to test for evidence to the contrary of this research - that experienced drivers already possess the skills to handle control transitions, and that this would be demonstrated given enough salient cues from the external environment (i.e. a higher level of traffic density and interaction borne from increased realism in the test environment). Therefore, in this final study, drivers in both experimental groups were provided a more realistic set of scenarios and traffic behaviour, and a multitude of salient cues within the external environment. The aim was to offer as much opportunity as possible, through bottom-up information, to instigate similar glance behaviour patterns across the groups. Despite this, drivers who did not receive a behavioural intervention demonstrated insufficient or inappropriate glance behaviour that resulted in a level of SA that was too low and workload that was too high to meet all levels of task demands from the driving environment (Fuller, 2000). This was evidenced by the lower levels of attention given by these drivers to essential perceptual information in comparison to the group who received the intervention. This subsequently lead to ill-timed off-road glances and critical information being missed. This mismatch between the capability of these drivers and the task demands of the driving environment, resulted poor safety outcomes, notably higher levels of near crash events (Seppelt et al., 2017) and incidents of cognitive tunnelling (Endsley, 2005).

Qualitative findings from the thematic analysis presented in chapter five highlight a deeper understanding of driver views regarding the necessity for additional support. For example, drivers, regardless of whether they had received the training invention, acknowledged that there was a need to take on responsibility for the monitoring task during the transition period. However, this was accompanied by a recognition and concern relating to the automatic processing of skill-based behaviours typically involved in the driving task, and the likelihood that (as demonstrated by those who had not received intervention) drivers would not conduct the behaviours required to rebuild attention and SA during the PCTP. Additionally, data from the interviews and simulator study suggested that drivers will likely glance at the forward roadway, but are less likely to take in the surrounding road. Instead, they are more likely to continue to demonstrate a high priority towards internal areas of interest. Findings suggest that the temporal glance patterns of these drivers towards the external environment will also likely be inconsistent and fleeting in nature, likely missing the motivation and purposeful goal of raising and recalibrating their SA prior to taking control of the driving task – something that is essential for anticipatory

response behaviour to potential hazards (Seppelt et al., 2017; Pipcorn et al., 2024).

These findings highlight that drivers found it difficult to return to their well-rehearsed and automatically undertaken glance strategies prior to re-engaging with the motor control loop following resumption of manual driving (Mole et al., 2019). In addition, glance behaviour analysis from the final driving simulator study demonstrated that this had a cumulative, detrimental impact on glance behaviour patterns following manual mode resumption (Pipcorn et al., 2024). Taken together, these results provide decisive evidence that drivers find it challenging, not only to remember or, in fact, recognise the importance of attention allocation to the external driving environment during the PCTP, but also to actively resume appropriate and effective scanning behaviour without additional training or instructional guidance (Endsley, 2017a; Carsten and Martens, 2019; Mole et al., 2019).

Themes from the interview data also offered useful insight into the drivers' mental models of L3 ADS and their associated expectations and understanding of driver responsibilities. Comments highlighted that, in line with the literature (McDonald et al., 2018; Casner and Hutchins, 2019), many drivers do not hold accurate mental models of L3 ADS, in particular, in relation to the concepts of shared control, joint activity and the required engagement with the driving task during a PCT and prior to resumption of manual driving. This was made even more salient by the controlled nature of PCTs, or, more specifically, the lack of urgency and criticality involved due to the rationale behind these ODD boundary driven requests for intervention. Outputs from the interviews suggested that driver's mental models of control transitions and their understanding of vehicle automation had been influenced by the high-profile nature of failure-induced control transitions and their potential consequences in the media. As drivers have witnessed examples of ADS capability in real-time, in particular in relation to OEDR (Banks et al., 2018; Lin et al., 2018; Brown and Laurier, 2019; Endsley, 2017a). Subsequently, the views and expectations relating to drivers' roles and responsibilities expressed during the interviews were linked to identifying hazards or critical situations that the vehicle is unable to handle so that the driver can immediately intervene.

There were a number of behavioural patterns and interview comments by drivers who did not receive any intervention that could be attributed, in part, to this understanding of the characteristics and capabilities of L3 ADS. Prior to covering these, it is worth noting that it is recognised that the SAE J3019 highlights that at level 3, the requirement for the driver continues to include fallback responsibility for failure induced control transitions. However, it is the assertion (and therefore assumption) of the approach taken by this thesis that, if the expectation related to L3 technology is that drivers can carry out 'eyes off road' activities thereby

relinquishing responsibility of the monitoring task to the ADS, then system reliability must be such that the majority of required interventions are due to approaching ODD boundaries. In which case, advanced warning can be provided to the driver along with sufficient time to rebuild SA prior to resuming the driving task (Trosterer et al., 2017; Carsten and Martens, 2019).

The underlying belief that control transitions are predominantly failure induced provides an explanation for the transient nature of the glance behaviour during the PCTP by drivers who did not receive an intervention. These drivers expressed during their interviews that when they received the TOR, they sought to find a salient and immediate hazard within the front roadway that caused the request to intervene. Consequently, the lack of any obvious cue led to complacency when what they observed did not match their expected requirement to attend to the road scene. Interestingly, the number of these drivers who did not spot the hazard car to the rear of the ego vehicle during this experimental scenario, suggests that the expectation of a critical incident does not automatically prompt the driver to check the surrounding roadway. These findings suggest that the expectation that takeovers are required, predominantly, for human intervention in critical situations, places a narrow, and reductive framework around the driving task in general, and subsequently, the driver's role and behavioural requirements during control transitions. Under this framework, the focal points for driver performance are firstly, how quickly can the driver respond to takeover request and secondly, how effectively can they respond to the hazard. This links the driver's responsibility requirements to the identification of a specific critical incident and associated hazards and the control response inputs to handle the incident.

There are a number of challenges associated with this approach. Firstly, there is no anticipatory requirement inherent in this role. The expectation on the driver is one of a reactionary response to a presented critical issue. In fact, part of the themes that came out of the driver interviews was that drivers expected and wanted the HMI to provide them with hazard detection information, so that they would be better equipped to respond faster and more effectively. In theory, this appears to be a sensible approach. However, in practice, and as demonstrated by Eriksson et al. (2018) and theorised by Endsley (2017), the use of information cueing instead of encouraging operators to use their own sense more effectively can lead to issues of cognitive tunnelling and the potential for an exponential increase in automation bias over time.

Secondly, the temporal requirement of visual behaviour inferred by this responsibility is rapid, fleeting and temporary in nature. A predominant focus on failure-induced control transitions addresses the urgency involved in dealing with a critical scenario. However, it does not take a

holistic view of the visual requirements associated with the ongoing management of the full driving task. The findings of this thesis highlight the importance of focusing on strategies that support the driver in the redistribution and ongoing responsibility of the full driving task, rather than focusing on handling an individual critical event, regardless of the reason behind the required intervention. As highlighted by Seppelt et al. (2017), research highlights drivers' ability to rapidly gain the 'gist' of a momentary scene, something needed in the face of needing to handle a critical incident. However, the visual requirements associated with anticipatory hazard identification and response are gained over a period of increased attentional allocation and the amalgamation of multiple dynamic scenes that enable high levels of SA. The focus on failure-induced control transitions within current research encourages strategies to improve the former and ignores ways to improve the latter: a gap this research has highlighted and contributed towards filling.

Finally, a lack of focus on PCTs ignores the standardisation of this requirement as new sub-task. The nature of a failure-induced control transition, particularly at level 3 if drivers are going to be permitted to carry out NDRTs, is that it is anticipated (and hoped!) that they won't happen. However, given the expected capability and regulatory operational design domain boundaries (e.g. geofencing to motorways) of L3 ADS, it is likely that PCTs will be a requirement for most, if not all, journeys. As highlighted throughout this thesis, it is likely that the assumption that experienced drivers already possess the skills to handle PCTs is partly responsible for the lack of focus on this sub-task or recognition that this presents a gap in driver competency. However, it was also the prediction of this research, as expressed in the design of the final simulator study, that for a large amount of PCTs this may represent the observed outcome. For example, the delay in the return of manual driving levels of visual scanning behaviour following a period of automated driving on a test track reported by Pipcorn et al. (2024) was mirrored in the detrimental effects on visual behaviour found in drivers who did not receive the intervention in the final simulator study.

As anticipated, the performance impact of reduced attention during and after the PCTP (as reported earlier in this discussion), was not observed until the driving task demand reached a level higher than the momentary capability available by drivers within this group. The resultant impact of the lack of instructional guidance on how best to carry out a PCT (i.e. a gap in driver competency), along with the increased human factors issues (i.e. reduced attention and SA) due to inappropriate or insufficient behaviours and actions of drivers during the PCTP, meant that their available capability following the transition to manual driving was limited. This was not an issue whilst the levels demanded by the driving task were equal to or lower than the driver's available capability. However, a spike in the task demand, as expected

within a dynamic environment that can change suddenly and unpredictably (Teh et al., 2014), resulted in examples of loss of control, such as increased driver error and near crash incidents.

The inconsistent and transient nature of visual behaviour during the PCTP by drivers who did not receive an intervention was exacerbated by an apparent dichotomous and delineated view relating to the responsibility for the driving task dependent on who had primary control of operations. This was reported in the way these drivers spoke about their approach to the takeover, and their expectations around the responsibility for the monitoring task during the transition period. It was also demonstrated, not only in drivers' visual behaviour, but also through the increased likelihood of these drivers to continue to engage in their NDRTs following the issue of a TOR. Higher levels of shared attention during the PCTP compared to drivers who received an intervention were found in the observation study and both simulator-based studies, including the final high task demand drive, despite the number of external cues. The method employed by all three of these studies (i.e. to allow drivers the freedom to use their own NDRTs as they saw appropriate) was of significant benefit here. Including personal devices and affording autonomy to the drivers increased the ecological validity of driver behaviours in the study. This subsequently facilitated the observation of potential driver operational deviations due to the motivational aspects that were introduced by not using a prescribed secondary task.

Whilst it could be concluded that an increase in NDRT use might be expected in a group with no intervention and that this problem could be resolved by restricting access to secondary devices following the TOR, the research findings presented in this thesis would suggest such an approach would be too simplistic, potentially detrimental to the psychological comfort of drivers with a hard cutoff to device access, and ignore the overarching need to impart understanding of the shared nature of this new driving task. It is the assertion of this research that the behaviour demonstrated by drivers who were not exposed to a behavioural intervention highlights the need for a human machine teaming approach to the system design and operation of vehicles equipped with L3 ADS and one which encourages joint cognition and partnership, in particular during safety critical tasks such as control transitions. In fact, results from the final simulator study suggest that, when presented with an intervention aimed to encourage a collaborative and joint approach to the driving task, drivers were more likely to engage (appropriately) with NDRTs during automated driving and more likely to discharge them earlier following a TOR than their counterparts who were left to decide how to behave during the same amount of time between the TOR and manual driving mode engagement.

Taken together, the evidence presented here offers additional context to explain observed driver behaviour and bolsters support for the need to expand the research approach into control transitions and the necessity for design solutions to improve driver competency in this new sub-task. These findings caution against taking a reductionist approach to resolve the challenges raised by the human operation of L3 ADS. An approach that focuses on tackling individual challenges through the use of, and reliance on even more automation, such as, hazard alert systems or entertainment systems that cut a driver off mid-activity, risks becoming a reactionary ‘whack-a-mole’ strategy to handle the multitude of potential issues this technology could introduce, whilst ignoring both the bigger picture and the human who sits at the centre of the system. Instead, this research recommends an approach that views the driving task from a holistic sense and focuses on behavioural change strategies that includes imparting a mental model of shared control and guiding the joint activity of the driving task with the automated systems.

The results of behavioural observations of driver interactions with L3 ADS, as well as information collated from an extensive review of the literature, in particular taking inspiration from strategies used within the aviation domain, and the themes extracted from driver interviews were used to inform the design of the initiatives developed during phase three of the thesis. This phase looked to explore how driver training or technological solutions, such as HMI system design, could support desired driver behaviour, and mitigate against undesirable behavioural adaptations during PCTs. An evaluative discussion of the effectiveness, practicality and sufficiency of these interventions is presented in the next section.

8.3 Evaluation of design solutions

Research from the aviation domain, alongside the analysis of driver behaviour during the initial observation study, inspired the concept of the Check, Assess, Takeover (CHAT) standard operating procedure, which, in turn, informed both the training and technological intervention designs. Research into a number of high-profile aviation crashes following the introduction of automation emphasised the use of standard operating procedures (SOPs) as an integral part of the activities required to ensure successful operations in complex human-machine systems. It was highlighted that their use provides a way to standardise and specify the manner (and order) that tasks should be carried out. These procedural frameworks provided a checklist of actions that could be rapidly recalled to ensure the optimal, logical and safe sequencing of tasks, through clear and predictable instruction to the human operator (Degani and Wiener, 1997).

In the automotive domain, the utility of an operational procedure has already been demonstrated with the use of the mirror, signal,

manoeuvre routine (DVSA, 2024). Therefore, the familiarity associated with this approach to instruction was viewed as an additional benefit and evidence for its potential as a design solution. Additionally, further research (Trosterer et al., 2017) in the context of L3 ADS, drew links between pilots' use of SOPs and the pilot flight display to suggest potential system design approaches to support and guide drivers undertaking control transitions. These findings, drawn from pilot interviews, provided further support for an automotive automation operating procedure to standardise the sequence of actions to carry out the planned control transition sub-task.

The 'CHAT' (Check, Assess, Takeover) SOP was designed to be applied following a TOR for a planned control transition (PCT). Its concept was based on a recognition that the basic skills involved in vehicular control are not novel to this group of drivers, but rather it is the procedural rules and sequential requirements involved in the application of behaviours and actions to prepare for the transition that are of importance. The 'check, assess, takeover' sequence was designed based on patterns of behaviour identified in the observation study, which suggested that drivers would not recognise the importance of commencing actions associated with situation monitoring prior to taking over the full driving task.

The CHAT sequence aimed to guide the driver to make appropriate checks and assessments of their internal and external environments prior to taking over operational control of the vehicle. Additionally, the semantic association between the word CHAT and the process of communicating with a partner was deliberate. It was designed to provide reinforcement of the mental model of shared control and joint activity within a team, and in particular, the collaborative approach required during any kind of handover/takeover within a team, where information such as system state and action requirements needs to be passed between agents to safely control and perform the transition. The overarching aim of CHAT was to provide drivers with a simple and efficient way of remembering a specific sequence of actions that they must take following receipt of a TOR, thereby encouraging the driver to transition to an 'on the loop' control state, and in so doing, increasing their situation monitoring and SA before physically resuming manual control of the vehicle. The CHAT procedure was applied within two different design solutions: a behavioural training module and an HMI display. These interventions were evaluated using simulator-based driving studies. Additionally, interviews conducted with drivers who took part in the training study provided further insight into the effectiveness and potential utility of the training module, alongside potential limitations of this approach.

The CHAT behavioural training was presented in two parts. The first part focused on providing drivers with more comprehensive,

explanatory knowledge for mental model formation, which aimed to increase driver's knowledge and understanding of automation. The second part involved the use of the CHAT standard operating procedure, which aimed to provide these drivers with a simple, clear and concise procedure that could be used to efficiently recall the sequence of tasks required during planned control transitions. Findings demonstrated improvements in the visual behaviour of drivers who undertook the behavioural training in comparison to those who were asked to read through a user manual based on genuine instructions provided by a manufacturer of a L2 ADS equipped vehicle but updated to reflect L3 ADS capabilities. Drivers in receipt of the CHAT training were more likely to discharge their NDRTs sooner; conduct more mirror checks both during the PCTP and in preparation for a required lane change manoeuvre; and spot a hazard to the rear of the ego vehicle presented just before the resumption of manual driving.

These results suggest a positive response to the behavioural training that lean towards a 'proof-of-concept' design. However, there were a number of limitations that bounded the purported success. Firstly, drivers were only tested once, therefore, the effectiveness of retention of training could not be evaluated. Secondly, the level of sophistication offered by the simulation software used to develop the simulated environment was limited. This will have not only affected the level of immersion and affective risk felt by the drivers (albeit there were a number of behavioural responses that suggested a high enough level of fidelity to draw conclusions from the results, for example, finishing a hot cup of coffee so that they could 'safely' take over), it will have also impacted the number of external cues available to drivers to robustly test both the effectiveness of the training intervention, but also the necessity for it. This point was highlighted within the interview data, where drivers in both groups noted that they felt the low traffic density and level of interaction with the ego vehicle was the reason for their low levels of attention towards the external environment.

It is important to view results relating to initiatives aimed to improve performance in a safety critical system with a critical eye. Therefore, a second question around the robustness of the training's effectiveness is one of priming effects. Priming is a psychological phenomenon denoted by the rapid and automatic deployment of visual attention towards relevant stimuli (Kritjansson and Asgeirsson, 2019). Priming is not necessarily problematic; after all, part of the aim of using a standard operating procedure is to prompt a rapid recall of actions required. However, to fully achieve the aim behind the CHAT concept, it needs to be accompanied by a shift in understanding of the explanatory knowledge that sits behind the behavioural instruction. Otherwise, any mirror glances undertaken risk being performative and linked to the

narrow action of looking in a mirror, rather than purposeful, and linked to the broader goal of increasing SA.

Priming, by its nature, is suggestive of superficial stimulus-response behaviour rather than a deeper level of cognitive processing or a true influence of top-down information (Horner and Henson, 2008). In other words, could the reason that more drivers in the behavioural group saw the hazard vehicle and made more mirror glances than those in the operational group, simply be that they had been primed to check their mirrors, rather than because of an updated mental model and improved understanding related to the broader goal of situational monitoring and raising situational awareness in response to the training? By the same token, the interpretation of the increase in mirror glances in preparation for the lane change could be evidence of the cumulative effects of increased visual attention and improved scanning strategies during the PCTP. Equally, it could be argued that this increase was simply a reflection of these drivers responding to the hazard they spotted, potentially due to the effects of priming.

The use of an adapted version of the critical decision method interview, which was employed to elicit a deeper, contextual understanding about the drivers' cognitive processes during the experimental drive, went some way to begin to answer this question. Positively, the thematic analysis of this data suggested that these drivers were not just subject to the effects of priming with qualitative data from drivers who undertook the behavioural CHAT training pointing to an immediate potential benefit of enhanced knowledge and subsequent influence on driver behaviour during the transition period. Driver comments from this group provided evidence that the training was effective in raising awareness for the necessity for glance behaviour that provided ongoing sampling of the external driving environment to support SA recalibration and decision making into the conscious thought processes of drivers. In contrast, drivers in the operational group demonstrated a continued focus on behaviour and actions relating to the operational control of the vehicle and taking back physical control as quickly as possible, notably without any recognition of the importance, or need for increased attention and vigilance with the goal of regaining SA, outside of seeking the hazard that caused the TOR.

Taken together, it can be concluded with some degree of confidence, that the quantitative evidence from the experimental drive, alongside the qualitative evidence from the interviews, demonstrates a level of effectiveness beyond one of simple priming effects, and points towards the potential of the CHAT concept and behavioural training approach. However, the limits within the experimental design in relation to, firstly, a lack of assessment of retention; and secondly, a limited ability to assess the impact of the training on driver SA and performance after manual mode resumption, were things that were focal within the design

of the final simulator study. Additionally, outside of the analysis of NDRT use and driver shared attention following the TOR, there was no method of analysis employed during the first simulator study to assess how driver behaviour changed across the course of the PCTP. This was viewed as something that would provide useful additional context to understand the sequential behaviour of drivers and cognitive decision making throughout the transition period and, therefore, built into the design of the final study.

Additionally, comments from drivers supported the view that there is a need to support and guide drivers to recognise the importance of attention allocation to the driving environment during the PCTP and remember to do it (Endsley, 2017a; Mole et al., 2019; Carsten and Martens, 2019). However, drivers in both groups raised concerns about the utility of a formal training approach: firstly, relating to retention of training, highlighting that a one-off training module may be inadequate as there is a need to provide guidance in real-time when drivers are performing a transition of control; and secondly, outlining, unsurprisingly, that formal, mandatory driver training approaches may not be appropriate for all types of drivers, for example experienced as opposed to novice drivers.

In-vehicle HMIs offer a way of providing standardisation in the undertaking of a task through the delivery of the CHAT SOP in a real-time and accessible way. Not only does this provide a flexible and pragmatic way of reaching experienced drivers with consistent instructional guidance that could be provided to novices through formal training routes, it also offers a way to reinforce learning and memory of the operating procedure across all groups of drivers. The final stage of the research in this thesis aimed to address these concerns, through the application of the CHAT SOP to the development of an HMI display design. The aim of this initiative was to demonstrate the potential breadth and utility available with the application of the CHAT concept, demonstrating the ability of the CHAT SOP to provide a tailored and layered approach to training and guidance that could suit multiple categories of drivers.

As was highlighted during the extended literature review, the use of HMI design has been flagged as a suitable platform for an initiative to support drivers in the uptake of visual behaviour associated with situation monitoring and the rebuilding of SA (e.g. Endsley, 2017a; Trosterer, 2017; Carsten and Martens, 2019). This would also match with similar solutions provided within the aviation domain (Trosterer et al., 2017) and, as with in-vehicle navigation systems, offer a pragmatic source of real-time support for drivers to influence and guide decision making and action. However, for design to be effective, it is important that the human is considered within its development. This consideration extends not only to what is communicated through the HMI, but also

how it is communicated. The CHAT HMI was developed, once again, from the basis of guiding the sequential action and behaviour of the drivers using the CHAT concept. However, it also aimed to evoke formation of a mental model based on joint and collaborative activity that would promote a common-ground understanding of the human-automation system. In order to achieve this second aim, the HMI design took influence from research into the use of naturally occurring speech and conversation user interfaces. Evidence of the potential effectiveness of this approach was found from the driver interview data. Comments from a number of drivers across both experimental groups highlighted their desire for additional support from the in-vehicle HMI and the use of speech and vocal feedback from the vehicle systems. This suggested further support, in accordance with the literature (e.g. Large et al., 2019; Clark et al., 2024), for the use of a human, anthropomorphic quality to the in-vehicle system to convey a team approach to required control transitions.

The evidence suggests the achievement of a key goal of the CHAT HMI procedural design (Degani and Weiner, 1997), namely, to standardise and solidify the timely and efficient scheduling of the behaviour and actions required for the sub-task of PCTs. Drivers who used the CHAT HMI demonstrated collaboration with the vocal advisements from the HMI, matching their behaviour to the sequential checklist of required action. This included the timely discharging of NDRTs and higher levels of sustained and temporally interspaced visual attention to the external driving environment (Seppelt et al., 2017), compared with drivers who received the same time period to ready themselves for the resumption of manual driving, but did not receive the CHAT proforma. Analysis of drivers' glance data provided a clear indication that drivers who used the CHAT HMI understood what was expected of them (Carsten and Martens, 2019) and that the HMI design afforded the appropriate use and interaction with vehicle automation (Merat et al., 2019).

Additionally, these drivers showed evidence of positive behavioural adaptation over time that suggested their learning had been reinforced. For example, in stark contrast to the other driving group, the CHAT HMI drivers' responses to discharging NDRTs and getting ready for driving became more efficient over the course of the experimental drives and their visual behaviour consistently matched the task demand required from the external environment. Notably, the CHAT HMI design appeared to provide psychological comfort to these drivers that facilitated the use (and importantly the appropriate use) of NDRTs, whereas comfort levels for drivers who were left to their own devices during the PCTP appeared to decline, as evidenced by the reduction in engagement with NDRTs. One of the key advantages of driving automation is for drivers to be able to maximise turning their attention to other tasks. Therefore, this observation would seem to point to another

benefit (and reason) for introducing a behavioural intervention that helps to build trust, understanding and expectation between the driver and the ADS.

Several elements of the final simulator study design allowed for increased robustness, validity and supporting evidence to be provided in support of the effectiveness of CHAT and its impact on driver SA and performance. The introduction of Fuller's (2000) Task Capability Interface Model also provided a theoretical framework to shine a light on the importance of PCTs, despite their non-urgent status, and add weight to the necessity for drivers to receive training and guidance to carry them out. The use of three drives of differing task demand allowed for driving performance to be assessed between the two groups, over time and with varying levels of difficulty. In turn, this allowed for the assessment of both the necessity and effectiveness of the behavioural intervention; under differing circumstances and requiring varying levels of cognitive effort and complexity of decision making. The set up of the high task demand allowed for tangible evidence to be gathered on driver's glance behaviour and attention towards critical objects within the environment. Events following resumption of manual driving allowed for the analysis of quantitative and qualitative evidence to be used to draw inferences between SA development during the PCTP and driver glance scanning behaviour and performance. These elements together presented a much clearer picture of the links between the effectiveness of the CHAT protocol, improved driver SA, and, most importantly, its implications for driver performance and safety.

8.4 Overall Conclusion

The thesis lays out a clear challenge to the assumption that passive training methods are sufficient to provide experienced drivers with the requisite guidance required to safely interact with vehicles equipped with L3 ADS technology. The implications of the evidence presented are that consideration needs to be made for the provision of instructional guidance that focuses on how drivers should behave and interact with their vehicle's ADS to effectively co-ordinate the distribution of responsibility and control of the driving task. In contrast, typical training methods for experienced drivers, such as user manuals or dealership guidance, are not only less likely to be accessed by drivers, but they also concentrate on the operational aspects of vehicle systems (e.g. how to turn the automation on and off).

Notwithstanding the achievements reported in this thesis; there are a number of limitations that should be acknowledged, most notably, the limited number of drivers tested, the time they spent using the ADS (De Winter et al., 2021), and the use of a driving simulator as a testing environment. Fundamentally, there are always limitations to take account of when time, money and resources are bounded and the

subject in question is disruptive technology that offers a degree of safety concerns that requires adjustments to be made so that no harm comes of unwitting study volunteers! However, it is the view of the author that the results presented across the course of this thesis would suggest enough evidence to warrant further development. And this view is supported by recent evidence on the reduction of visual attention from driving studies using a test track that highlight the neglect and importance of driver's visual behaviour following automated driving (Pipcorn et al., 2024).

Additionally, it is worth noting that the scenarios presented here were not out of the ordinary and represent common situations that drivers could face coming out of an ODD boundary-provoked control transition. It is the assertion of this thesis that these elements caution against a continued sole focus on urgent takeover scenarios and, therefore, takeover time and hazard response. It is the belief of the author that the potential issues relating to available driver capability following a planned control transition due to a gap in driver competence is both obscured and downplayed due to the focus on critical (and often edge-case) scenarios related to failure-induced control transitions. It is suggested that the assumption that if the control transition is planned, drivers will be ready to take on the continued driving task, does not sufficiently consider the impact of automated driving on drivers' visual behaviour or recognition of the sudden, dynamic and unpredictable nature of the driving task that may challenge the driver shortly after they have resumed manual driving. It is suggested that future research should increase the focus and inclusion on both PCTs and driver visual behaviour to ensure a holistic view is taken of the driving task and its impact on driver safety and required performance.

8.5 Future Work and Design development

Firstly, on the subject of training necessity. The proposal made by the thesis is that planned control transitions represent a new sub-task within the overall task of driving. Within the context of driving automation, this sub-task requires a particular course of action from drivers that is not required within manual driving. Therefore, experienced drivers will not have any specific, formal instruction, training or experience to draw upon to demonstrate competence in this sub-task. The evidence presented by this research suggests that observed behavioural adaptations of experienced drivers undertaking PCTs without guidance reduces the available driver capability to handle sudden and dynamic changes in driving task demand following resumption of manual driving. It is recommended that further empirical work is carried out to increase the validity of these findings; for example, undertaking driving-based simulator studies with a significantly larger population of drivers; extended time within automated driving to review the impact on driver capability after takeover; or test-track based studies to review drivers

outside of the simulator. It would also be beneficial to review the behavioural adaptation of drivers over a longer period of time with repetitive use of these systems. However, it would likely only be possible to get a true sense of the impact of these systems once they are already available for public consumption, at which point, it becomes much more difficult (if not impossible) to address issues of safety and performance.

It is recognised that there are many alternative ways to approach the safety and performance aspects of control transitions within the driving domain, many of which could involve heightened regulatory restrictions relating to the ODD of L3 ADS. For example, at the time of writing, the current UK proposal related to the use of L3 ADS is both geo-fenced (motorway network only) and speed limited ($\leq 60\text{kph}/37.28\text{mph}$). There is also continued discussion about how, or if, drivers will be able to engage with secondary tasks, while the ADS is active. Additionally, system design approaches to counteract and provide redundancy against the effects of adverse operational deviations by human drivers are being explored, for example, driver monitoring systems, to encourage drivers to maintain a (physical) readiness to potential transition requests, and ADS controlled minimal risk manoeuvres in the event of a driver failure to respond to a TOR (Kinnear et al., 2020). The evidence regarding the potential challenges relating to the decrement of driver capability in both the short term (i.e. through driver disengagement during automated driving) and long term (i.e. driver skill degradation) suggest that other areas of safety should certainly continue to be considered. For example, driving research highlights factors within the driving environment that increase task demand (e.g. traffic density, vehicle position etc.). Therefore, a further area of recommended research would be to explore when and how PCT TORs should be issued by the ADS to ensure the task demand is as low as possible during and following the transition period. However, there is only so much control that can be applied to a dynamic and unpredictable environment, or indeed, the human driver. Therefore, the view of this thesis is that regardless of the level of control and structure placed around the system, it does not negate the need to provide updated instruction and guidance to the human operators using it. Future work on how, when and where L3 ADS should be used, must include the transfer of knowledge and understanding to the human agent, regarding the co-operative, behavioural requirements to competently undertake the driving task in this new context.

Secondly, the evidence from the simulator studies and thematic analysis presented in phase 3 of the thesis, suggests that the CHAT SOP has the potential to provide an effective structure for the application of standardised, guiding principles, behaviour and action for undertaking PCTs. However, the experimental setup of the driving

simulator used to evaluate the CHAT design solutions either limited or completely removed the authority available to drivers to actively take back control from the vehicle. The rationale for doing this was that it allowed for a controlled comparison of driver behaviour during the PCTP across groups. This was particularly important during the final study, where a comparison was being made between two groups of drivers who were provided different HMI designs – providing all drivers with the same amount of time to raise their SA and compare behaviour and performance was an important feature of this design. Future work should consider how potential design solutions could be assessed whilst giving the driver full control over the timing of the transition to manual control. Of particular note is in the driver training simulator study, where drivers were given some authority to take back control earlier than the full minute (albeit, this would still give the drivers ten seconds rather than jumping to full manual control). Drivers who did not receive CHAT training tended to activate this option without having conducted adequate glances to the surrounding road. Some drivers even activated this option when their eyes were not on the road at all. System design that gets the right balance between the conflicting issues of driver competence and driver authority to enact a change in the distribution of responsibility and control of the driving task (Flemisch et al., 2012) is a key challenge that warrants further exploration.

HMI design solutions offer a particularly pragmatic and flexible approach to support and guide experienced drivers in the undertaking of PCTs and the broader development of appropriate mental models related to L3 ADS. The evaluation of the CHAT HMI in its current form offers some positive initial evidence that demonstrates the potential utility and effectiveness of this medium from two perspectives. Firstly, in the delivery of SOP-style guidelines to drivers and the benefits related to the temporal order and rhythm of a behavioural checklist delivered in this manner. Secondly, in the use of conversational user interface design to encourage co-operative interaction and shared control in a way that provides psychological comfort to drivers through the delivery of instruction in a style that draws from natural human-human communication. These findings provide weight to justify future work to extend the proof-of-concept design presented here.

A key area recommended for future development of the CHAT HMI is the introduction of an interactive element within the design. Taking further inspiration from the research by Clark et al. (2024) that influenced the initial iteration of the CHAT HMI, it is suggested that the findings from their analysis of human-human communication during PCTs could be used to guide the structure of a more interactive style of conversational user interface. For example, their analysis of the communication between participants taking on the roles of either the driver or ADS, highlighted a number of interaction types used by

participants playing the role of the human driver during a takeover scenario. Further analysis revealed that certain interaction types, identified when participants took on the role of the ADS, commonly paired with certain driver interaction types. Dialogue categorised under the driver interaction type, 'the co-ordinator' was identified as typically pairing with the ADS interaction type 'the supervisor', which was used to develop the audio speech cues used in the current CHAT HMI design. Additionally, 'co-ordinator' dialogue is described as focusing on acknowledgements and disclosures that relays preparedness to receive control that would allow drivers to confirm receipt and acknowledgement of instructions from the HMI. Subsequently, this interaction may provide a way of introducing checks into the system that could balance the authority of drivers to take control of the driving task with confirmation that they have built up the requisite ability to redistribute responsibility and control from the ADS (Flemisch et al. 2012).

The final recommendation for future work also relates to the introduction of an interactive element within the CHAT HMI system design, but from a subtly different perspective. It takes further inspiration from the aviation domain and is made in recognition of research outlining the challenges for human drivers in the calibration of the perceptual-motor loop following a control transition (e.g. Mole et al., 2019). In two of the experimental drives used in the CHAT HMI evaluation study, a lane change was required by drivers shortly after the transition in order to take the motorway exit. Drivers who were given the CHAT HMI design for this study demonstrated better performance following the transition of control and a lower number of near crash incidents. However, in general, they did not demonstrate better operational control in the performance of the lane change itself. This is not necessarily surprising given what is reported in terms of the challenge in maintaining perceptual motor mappings when the driver has been disengaged from both the monitoring and motor control loops of the driving task (Mole et al., 2019). Therefore, even with guidance to encourage the recalibration of SA and attention during the PCTP, it could be beneficial to explore ways to either remove the requirement for a lane change following MDM resumption and/or provide drivers with additional support to allow them to better co-ordinate gaze patterns and control input.

The future design proposal is based on a procedure typically used by flight instructors in the management of dual control transitions with student pilots in a single aircraft (e.g. CAA, 2024). This procedure lays out the verbal commands and their paired actions, required to ensure all agents within the cockpit are aware of who is physically flying the aircraft. To initiate a control transition from the flight instructor to a student, standard aviation practice uses a follow-me-through exercise. Use of this exercise allows the student pilot to gain co-ordinated visual

and motor perception from the feel of control inputs that are initiated by the flying instructor. A verbal 'follow me through' instruction from the flight instructor initiates the control transition. The action required from the student pilot, following this instruction, is to lightly place their hands and feet on the controls and feel what the instructor is doing. Importantly, it is the flight instructor at this point who retains control of the aircraft. Following this action, the flight instructor will state, 'you have control' to the student pilot. This instruction requires an acknowledgment response from the student pilot, both verbally and physically. The verbal response is an 'I have control' statement back to the flight instructor and the physical response is a gentle lateral input to the control column to cause a slight movement of the aircraft along the longitudinal axis (Roll). Following the verbal and physical confirmation of preparedness to receive control from the student pilot, the flight instructor will state a final 'you have control' and will transfer control of the aircraft to the student pilot. The integration of this type of procedure into the design of the CHAT HMI offers the opportunity for further co-operation and co-ordination between the ADS and the human driver. For example, in a scenario where the vehicle is in the central lane when the TOR is issued, a 'follow-me-through' type procedure potentially offers an opportunity to both assist the co-ordination of driver gaze behaviour and control inputs with the ADS remaining in control, allows further clarity of control responsibility and confirmation of ability prior to completing the control transition, and reduces the task demand for the human driver (by retaining control of the lane change with the ADS) prior to transition to manual driving mode.

8.6 Closing Statement

From the evidence presented throughout this thesis, it is clear that the challenge raised against the assumption that experienced drivers do not require additional training or guidance to driver vehicles with L3 ADS carries weight. On the strength of the findings demonstrating the behavioural adaptations of drivers left to their own devices, it is argued that the provision of active forms of training or instructional guidance to drivers across the full spectrum of experience, is a necessity.

The specific competence gap identified and outlined by the thesis relates to the sub-task of planned control transitions. More broadly, it was highlighted that driver mental models need updating to include an understanding of the shared nature of the driving task and, subsequently, the co-operative expectations and responsibilities of the driver. The thesis recognises that intervention solutions put forward to close this competence gap need to provide consistency and standardisation across the driver population. However, it needs to be in a format that is simple and flexible to allow for application in both formal and informal instructional settings. The proposal of the thesis is that the CHAT standard operating procedure offers both of these elements. It

has demonstrated its potential utility as both a formal training module and a HMI design. The results of the evaluations of these design solutions demonstrate the potential of the CHAT procedure as an effective way of guiding desired driver behaviour. Future work has been recommended to both increase the validation of these findings and develop the concept of the CHAT design solutions to a required level of sufficiency to support driver performance and safety with L3 ADS.

References

- Abraham, H., Lee, C., Brady, S., Fitzgerald, C., Mehler, B., Reimer, B., & Coughlin, J. F. (2017). Autonomous Vehicles and Alternatives to Driving: Trust, Preferences, and Effects of Age Learning to Use Technology View project. *Transportation Research Board 96th Annual Meeting*, (January). Available at: <https://www.researchgate.net/publication/319269855>.
- Adavikottu, A., & Velaga, N. R. (2024). Modeling the impact of driving aggression on lane change performance measures: steering compensatory behavior, lane change execution duration and crash probability. *Transportation research part F: traffic psychology and behaviour*, 103, 526-553.
- Bagdadi, O. (2013). Assessing safety critical braking events in naturalistic driving studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 16, 117-126.
- Bainbridge, L. (1982). Ironies of Automation. *IFAC Proceedings Volumes*. Elsevier, 15(6):129–135.
- Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics*. Elsevier, 68:138–145.
- Banks, V. A., Plant, K. L. and Stanton, N. A. (2019). Driving aviation forward; contrasting driving automation and aviation automation. *Theoretical Issues in Ergonomics Science*. Taylor & Francis, 20(3):250–264.
- Banks, V. A. and Stanton, N. A. (2016). Keep the driver in control: Automating automobiles of the future. *Applied Ergonomics*. Elsevier Ltd, 53:389–395.
- Banks, V. A. and Stanton, N. A. (2019). Analysis of driver roles: modelling the changing role of the driver in automated driving systems using EAST. *Theoretical Issues in Ergonomics Science*. Taylor & Francis, 20(3):284–300.
- 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*. Elsevier, 62:505–514.
- 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: Traffic Psychology and Behaviour*. Elsevier Ltd, 35:75–84.
- 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*. Elsevier Ltd, 18:47–57.

Bell, B. S. & Kozlowski, S. W. J. (2008). Active Learning: Effects of Core Training Design Elements on Self-Regulatory Processes, Learning, and Adaptability Regulatory Processes, Learning, and Adaptability. *Journal of Applied Psychology*. 93(2):296.

Braun, V. & Clark, V. (2006). *Using thematic analysis in psychology, Qualitative Research in Psychology*, 3(2):77–101.

Brookhuis, K. A., de Vries, G., & De Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis & Prevention*, 23(4), 309-316.

Brown, J. P. (2016). The effect of automation on human factors in aviation. *The Journal of Instrumentation, Automation and Systems*, 3(2), 31-46.

Brown, B. & Laurier, E. (2017). The Trouble with Autopilots, in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI, 17. New York, New York, USA: ACM Press :416–429.

Budiu, R. (2018). Between-Subjects vs. Within Subjects Study Design. Available at: <https://www.nngroup.com/articles/between-within-subjects/> (Accessed 18 August 2020).

Burnett, G., Large, D. R. & Salanitri, D. (2019). *How will drivers interact with vehicles of the future?* RAC Foundation, London.

Caird, J. K., & Horrey, W. J. (2011). Twelve practical and useful questions about driving simulation. In D. L. Fisher, M. Rizzo, J. K. Caird, & J. D. Lee (Eds). *Handbook of driving simulation for engineering, medicine, and psychology* 5:1-5,18. Boca Raton, FL: CRC Press.

Campbell Darryl (2019). The many human errors that brought down the Boeing 737 Max - The Verge, Vox Media. Available at: <https://www.theverge.com/2019/5/2/18518176/boeing-737-max-crash-problems-human-error-mcas-faa> (Accessed: 18 September 2019).

Carsten, O. & Martens, M. H. (2019). How can humans understand their automated cars? HMI principles, problems and solutions, *Cognition, Technology and Work*. Springer London, 21(1):3–20.

Casner, S. M. & Hutchins, E. L. (2019). What Do We Tell the Drivers? Toward Minimum Driver Training Standards for Partially Automated Cars. *Journal of Cognitive Engineering and Decision Making*. 13(2):55–66.

Casner, S. M., Hutchins, E. L. & Norman, D. (2016). The challenges of partially automated driving. *Communications of the ACM*, 59(5):70–77.

Castro, C., Ventsislavova, P., Pena-Suarez, E., Gugliotta, A., Garcia-Fernandez, P., Eisman, E., & Crundall, D. (2016). Proactive Listening to a Training Commentary improves hazard prediction. *Safety Science*. Elsevier, 82: 144–154.

Civil Aviation Authority of New Zealand (CAA)(2024). *Flight Instructor Guide*. New Zealand Government. Available at: <https://www.aviation.govt.nz/assets/licensing-and-certification/flight-instructor-guide/CAA-Flight-Instructor-Guide-web.pdf>

Clark, J. R., Large, D. R., Shaw, E., Nichele, E., Trigo, M. J. G., Fischer, J. E., & Stanton, N. A. (2024). Identifying interaction types and functionality for automated vehicle virtual assistants: An exploratory study using speech acts cluster analysis. *Applied Ergonomics*, 114, 104152.

Crandall, B., Klein, G., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Boston Review.

Crouch, B. (2023). The State of Self-Driving Cars: Autonomous Advances. Where is Self-Driving Technology in 2023? – Techspot, TechSpot, inc.. Available at: <https://www.techspot.com/article/2644-the-state-of-self-driving-cars/#:~:text=Today%2C%20car%20manufacturers%20are%20racing%20to%20develop%20the,how%20far%20we%20still%20have%20yet%20to%20go>. (Accessed: August 2023)

Cummings, M., Huang, L., Zhu, H., Finkelstein, D., & Wei, R. (2019). The Impact of Increasing Autonomy on Training Requirements in a UAV Supervisory Control Task. *Journal of Cognitive Engineering and Decision Making*. Sage CA: Los Angeles, CA. 295-309.

Cummings, M. M. (2014). Man versus Machine or Man + Machine? *IEEE Intelligent Systems*, 29(5):62–69.

Cummings, M. (2018). Informing Autonomous System Design Through the Lens of Skill-, Rule-, and Knowledge-Based Behaviors. *Journal of Cognitive Engineering and Decision Making*, 12, 1, 58–61.

De Waard, D., Kruizinga, A., & Brookhuis, K. A. (2008). The consequences of an increase in heavy goods vehicles for passenger car drivers' mental workload and behaviour: A simulator study. *Accident Analysis & Prevention*, 40(2), 818-828.

De Winter, J.C., Eisma, Y.B., Cabrall, C.D.D., Hancock, P.A. & Stanton, N.A., (2019). Situation awareness based on eye movements in relation to the task environment. *Cognition, Technology & Work*, 21(1):99-111.

De Winter, J.C., Happee, R., Martens, M.H., & Stanton, N.A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence.

Transportation Research Part F: Traffic Psychology and Behaviour. Elsevier Ltd, 27(PB):196–217.

De Winter, J., Stanton, N., & Eisma, Y. B. (2021). Is the take-over paradigm a mere convenience?. *Transportation Research Interdisciplinary Perspectives*, 10, 100370.

Degani, A.; Weiner, E. (1994). Philosophy, policies, procedures, and practices in *Aviation Psychology in Practice*. Brookfield; VT: Ashgate Publishing Co.

Degani, A. & Wiener, E. L. (1997). Procedures in complex systems: the airline cockpit. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 27(3):302–312.

Dekker, S. & Woods, D. (2002). MABA-MABA or Abracadabra? Progress on Human-Automation Co-ordination. *Cognition, Technology & Work*, 4, 4: 240–244.

Department for Transport (DfT) (2020). *THINK! – Road safety*. Crown copyright. Available at: <https://www.think.gov.uk/> (Accessed: 26 June 2020).

Department for Transport (DfT) (2025). *Regaining Situational Awareness as a User in Charge: Responding to transition demands in automated vehicles*. Crown copyright. Available at: <https://assets.publishing.service.gov.uk/media/677fb7c98ef66f3f5ea3984b/regaining-situational-awareness-as-a-user-in-charge-responding-to-transition-demands-in-automated-vehicles.pdf> (Accessed: 25 Jan 2025)

Dogan, E., Rahal, M. C., Deborne, R., Delhomme, P., Kemeny, A., & Perrin, J. (2017). Transition of control in a partially automated vehicle: Effects of anticipation and non-driving-related task involvement. *Transportation Research Part F: Traffic Psychology and Behaviour*. Elsevier Ltd, 46:205–215.

Driver & Vehicle Standards Agency (DVSA)(2024). *Safer Driving for Life – the Mirrors – Signal – Manoeuvre (MSM) routine*. Crown copyright. Available at: <https://www.safedrivingforlife.info/advice/car-driving/develop-practical-driving-skills/car-driving-junctions/#:~:text=Always%20use%20the%20MSM%2FPSL%20routine%20at%20junctions.%20Mirrors%3A,in%20good%20time.%20Speed%3A%20adjust%20it%20as%20necessary.>

Endsley, M. R. (2017a). From Here to Autonomy: Lessons Learned from Human-Automation Research. *Human Factors*, 59(1):5–27.

Endsley, M. R. (2017b). Autonomous Driving Systems: A Preliminary Naturalistic Study of the Tesla Model S. *Journal of Cognitive Engineering and Decision Making*, 11, 3, 225–238.

- Endsley, M. R. (1988). Situation awareness global assessment technique (SAGAT). *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference*. IEEE: 789–795.
- Endsley, M. R. (2018). Situation awareness in future autonomous vehicles: Beware of the unexpected. In *Congress of the International Ergonomics Association* (pp. 303-309).
- Endsley, M. R. (2015). Situation awareness: operationally necessary and scientifically grounded. *Cognition, Technology & Work*, 17, 163-167.
- Eriksson, A. & Stanton, N.A., (2017a). Takeover time in highly automated vehicles: noncritical transitions to and from manual control. *Human Factors*, 59(4):689-705.
- Eriksson, A., & Stanton, N. A. (2017b). The chatty co-driver: A linguistics approach applying lessons learnt from aviation incidents. *Safety Science*, 99, 94-101.
- Eriksson, A., Petermeijer, S. M., Zimmermann, M., De Winter, J. C., Bengler, K. J., & Stanton, N. A. (2018). Rolling out the red (and green) carpet: supporting driver decision making in automation-to-manual transitions. *IEEE Transactions on Human-Machine Systems*, 49(1), 20-31.
- Eriksson, A., Petermeijer, S. M., Zimmermann, M., De Winter, J. C., Bengler, K. J., & Stanton, N. A. (2018). Rolling out the red (and green) carpet: supporting driver decision making in automation-to-manual transitions. *IEEE Transactions on Human-Machine Systems*, 49(1), 20-31.
- Field, A. (2018). GLM 5: Mixed designs. *Discovering Statistics Using IBM SPSS 5th Edition*. SAGE Publications, 703-734.
- Flemisch, F. et al. (2012). Towards a dynamic balance between humans and automation: Authority, ability, responsibility and control in shared and cooperative control situations. *Cognition, Technology and Work*, 14(1):3–18.
- Friard, O. & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*. British Ecological Society, 7(11):1325–1330.
- Fuller, R. (2000). The task-capability interface model of the driving process. *Recherche-Transports-Sécurité*, 66, 47-57.
- Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis & Prevention*, 37(3), 461-472.

- Fylan, F. (2017). *Using Behaviour Change Techniques: Guidance for the road safety community*. London.
- Gold, C., Berisha, I. & Bengler, K. (2015a). Utilization of Drivetime – Performing Non-Driving Related Tasks While Driving Highly Automated. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Human Factors and Ergonomics Society Inc., 59(1):1666–1670.
- Gold, C., Körber, M., Hohenberger, C., Lechner, D. and Bengler, K. (2015b). Trust in automation –Before and after the experience of take-over scenarios in a highly automated vehicle. *Procedia Manufacturing*, 3:3025-3032.
- Gold, C., Happee, R. & Bengler, K. (2018). Modeling takeover performance in level 3 conditionally automated vehicles. *Accident Analysis and Prevention*. Elsevier, 116(April 2017):3–13.
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*. North-Holland, 52(C):139–183.
- Haslbeck, A., & Hoermann, H. J. (2016). Flying the needles: Flight deck automation erodes fine-motor flying skills among airline pilots. *Human Factors*, 58(4), 533-545.
- Hatakka, M. et al. (2002). From control of the vehicle to personal self-control; broadening the perspectives to driver education. *Transportation Research Part F: Traffic Psychology and Behaviour*. Elsevier Ltd, 5(3):201–215.
- Herslund, M. B. & Jørgensen, N. O. (2003). Looked-but-failed-to-see-errors in traffic. *Accident Analysis and Prevention*. Elsevier Ltd, 35(6):885–891.
- Horner, A. J., & Henson, R. N. (2008). Priming, response learning and repetition suppression. *Neuropsychologia*, 46(7), 1979-1991.
- Janssen, C. P., Iqbal, S. T., Kun, A. L. & Donker, S.F. (2019). Interrupted by my car? Implications of interruption and interleaving research for automated vehicles. *International Journal of Human-Computer Studies*. Academic Press, 130:221–233.
- Johnson, M., Bradshaw, J. M., Hoffman, R. R., Feltovich, P. J., & Woods, D. D. (2014). Seven cardinal virtues of human-machine teamwork: examples from the DARPA robotic challenge. *IEEE Intelligent Systems*, 29(6), 74-80.
- Kaber, D. B. (2018). Issues in Human–Automation Interaction Modeling: Presumptive Aspects of Frameworks of Types and Levels of Automation. *Journal of Cognitive Engineering and Decision Making*. SAGE Publications, Sage CA: Los Angeles, CA, 12(1) 7–24.

- Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3), 203-220.
- Kinnear, N., Stuttard, N., Hynd, D., Helman, S., & Edwards, M. (2020). *Safe performance of other activities in conditionally automated vehicles: Automated Lane Keeping System* (No. PPR 979).
- Kircher, K., Larsson, A. & Hultgren, J. A. (2014). Tactical driving behavior with different levels of automation. *IEEE Transactions on Intelligent Transportation Systems*. IEEE, 15(1):158–167.
- Klauer, S. G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., & Ramsey, D.J. (2006). *The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data*. United States. National Highway Traffic Safety Administration. Available at: <https://vtechworks.lib.vt.edu/handle/10919/55090> (Accessed: 26 June 2020).
- Klein, G., Feltovich, P. J., Bradshaw, J. M., & Woods, D. D. (2005). Common ground and coordination in joint activity. *Organizational Simulation*, 53, 139-184.
- Klien, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems*, 19(6), 91-95.
- Körber, M., Baseler, E. & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, 66:18–31.
- Krampell, M., Solís-Marcos, I. & Hjälm Dahl, M. (2020). Driving automation state-of-mind: Using training to instigate rapid mental model development. *Applied Ergonomics*, 83(October 2019).
- Kristjánsson, Á., & Ásgeirsson, Á. G. (2019). Attentional priming: recent insights and current controversies. *Current Opinion in Psychology*, 29, 71-75.
- Kyriakidis, M., De Winter, J.C., Stanton, N., Bellet, T., Van Arem, B., Brookhuis, Martens, M.H., Bengler, K., Andersson, J., Merat, N., & Reed, N. (2019). A human factors perspective on automated driving. *Theoretical Issues in Ergonomics Science*. Taylor & Francis, 20(3):223–249.
- Large, D.R., Banks, V., Burnett, G. (2017a). Literature Review: Human-Automation Coordination, Communication and Shared Control. Nottingham.
- Large, D. R., Clark, L., Quandt, A., Burnett, G., & Skrypchuk, L. (2017b). Steering the conversation: A linguistic exploration of natural

- language interactions with a digital assistant during simulated driving. *Applied Ergonomics*, 63, 53-61.
- Large, D.R., Burnett, G., Salanitri, D., Lawson, A., Box, E. (2019). A Longitudinal Simulator Study to Explore Drivers' Behaviour in Level 3 Automated Vehicles in *Automotive User Interfaces and Interactive Vehicular Applications* (AutoUI2019). Utrecht, Netherlands.
- Lampert, A., Dale, R., & Paris, C. (2006, November). Classifying speech acts using verbal response modes. In *Proceedings of the Australasian Language Technology Workshop 2006* (pp. 34-41).
- Lawson, G., Shaw, E., Roper, T., Nilsson, T., Bajorunaite, L., & Batool, A. (2019). *Immersive virtual worlds: Multi-sensory virtual environments for health and safety training*. Available at: <http://arxiv.org/abs/1910.04697> (Accessed: 26 June 2020).
- Lee, J., Young, K. & Regan, M. (2009). Defining driver distraction. In Regan, M., Lee, J., & Young, K. (eds) *Driver Distraction: Theory, Effects, and Mitigation*. Boca Raton: CRC:31-40.
- Li, N. & Busso, C. (2013). Driver Mirror-Checking Action Detection Using Multi-Modal Signals. In *Proceedings of 6th Biennial Workshop in Digital Signal Processing for In-Vehicle Systems and Safety*, 101-108.
- Lin, B. T., & Bao, S. (2019, June). The effect of turn signal onset on lateral performance measures when overtaking a lead vehicle-using naturalistic driving environment. In *Driving Assessment Conference* (Vol. 10, No. 2019). University of Iowa
- Lin, R., Ma, L. & Zhang, W. (2018). An interview study exploring Tesla drivers' behavioural adaptation. *Applied Ergonomics*. Elsevier, 72 (December):37-47.
- Mack, A. & Rock, I. (1998). Inattention Blindness: Perception without attention', in Wright, R. D. (ed.) *Visual Attention*. Oxford University Press: 55-76.
- McDonald, A; Carney, C.; McGehee, D. V. (2018). *Vehicle Owners' Experiences with and Reactions to Advanced Driver Assistance Systems*. Washington, DC. Available at: <https://aaaafoundation.org/vehicle-owners-experiences-reactions-advanced-driver-assistance-systems/>.
- Melcher, V., Rauh, S. Deiderichs, F., Widloither, H., & Bauer, W. (2015). Takeover Requests for Automated Driving. *Procedia Manufacturing*, 3:2867-2873.
- Merat, N., Jamson, A. H., Lai, F. C., & Carsten, O. (2012). Highly automated driving, secondary task performance, and driver state. *Human Factors*, 54(5), 762-771.

- Merat, N., Jamson, A.H., Lai, F.C., Daly, M., & Carsten, O.M. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*. Elsevier Ltd, 27(PB):274–282.
- Merat, N., Seppelt, B., Louw, T., Engstrom, J., Lee, J.D., Johansson, E., Green, C.A., Katazaki, S., Monk, C., Itoh, M., & McGeehee, D. (2019). The “Out-of-the-Loop” concept in automated driving: proposed definition, measures and implications. *Cognition, Technology and Work*. Springer London, 21(1):87–98.
- Michon, J. A. (1985). A Critical View of Driver Behaviour Models: What Do We Know, What Should We Do? *Human Behaviour and Traffic Safety*. Edited by R. Evans, Leonard; Schwing. Boston, MA: Springer.
- Mole, C. D., Lappi, O., Giles, O., Markkula, G., Mars, F., & Wilkie, R. M. (2019). Getting back into the loop: the perceptual-motor determinants of successful transitions out of automated driving. *Human Factors*, 61(7), 1037-1065.
- Mosier, K. L., & Skitka, L. J. (2018). Human decision makers and automated decision aids: Made for each other?. In *Automation and Human Performance* (pp. 201-220). CRC Press.
- National Transportation Safety Board (NTSB)(2020). *Final Reports for 2 Advanced Driver Assistance System Crash Investigations* Published. [Press Release]. 19 March. Available at: <https://www.nts.gov/news/press-release/Pages/NR20200319.aspx>. (Accessed: 29/07/2020).
- National Transportation Safety Board (NTSB)(2019). *Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian*. Available at: <https://www.nts.gov/investigations/AccidentReports/Reports/HAR1903.pdf> (Accessed: 29/07/2020)
- National Transportation Safety Board (NTSB)(2017). Collision Between a Car Operating with Automated Vehicle Control Systems and a Tractor-Semitrailer Truck, *Highway Accident Report*. Available at: <https://www.nts.gov/investigations/AccidentReports/Reports/HAR1702.pdf>. (Accessed: 25/04/2019)
- Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books.
- Nylen, A. B., Reyes, M.L., Roe, C.A., & McGeehee, D.V. (2019). Impacts on Driver Perceptions in Initial Exposure to ADAS Technologies. *Transportation Research Record: Journal of the Transportation Research Board*:354-360.

Östlund, J., Nilsson, L., Carsten, O., Merat, N., Jamson, H., Jamson, S., & Brouwer, R. (2004). HASTE deliverable 2: HMI and safety-related driver performance.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced 'complacency'. *The International Journal of Aviation Psychology*, 3(1), 1-23.

Parasuraman, R. & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. SAGE Publications. Sage CA: Los Angeles, CA, 39(2):230–253.

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(3), 286-297.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51, 195-203.

Petermeijer, S. M., Abbink, D. A., & De Winter, J. C. (2015). Should drivers be operating within an automation-free bandwidth? Evaluating haptic steering support systems with different levels of authority. *Human Factors*, 57(1), 5-20.

Pipkorn, L., Dozza, M., & Tivesten, E. (2024). Driver visual attention before and after take-over requests during automated driving on public roads. *Human Factors*, 66(2), 336-347.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, SMC*, 13(3):257–266.

Rasmussen, J. & Vicente, K. J. (1989). Coping with human errors through system design: implications for ecological interface design. *International Journal of Man-Machine Studies. Academic Press*, 31(5):517–534.

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, 39(3), 490-500.

Sarter, N. B. & Woods, D. D. (1994). Pilot Interaction With Cockpit Automation II: An Experimental Study of Pilots' Model and Awareness of the Flight Management System. *The International Journal of Aviation Psychology*. Lawrence Erlbaum Associates, Inc., 4(1):1–28.

Searle, J. R., Kiefer, F., & Bierwisch, M. (Eds.). (1980). *Speech Act Theory and Pragmatics* (Vol. 10). Dordrecht: D. Reidel.

Seppelt, B. D., Seaman, S., Lee, J., Angell, L. S., Mehler, B., & Reimer, B. (2017). Glass half-full: On-road glance metrics differentiate crashes from near-crashes in the 100-Car data. *Accident Analysis & Prevention*, 107, 48-62.

Seppelt, B. D. & Victor, T. W. (2016). Potential Solutions to Human Factors Challenges in Road Vehicle Automation. *Springer, Cham*:131–148.

Shaw, E., Large, D. R. & Burnett, G. (2020a). Towards Future Driver Training: Analysing Human Behaviour in Level 3 Automated Cars. In Charles, Rebecca; Golightly, D. (ed.) *Contemporary Ergonomics & Human Factors*: 59–66.

Shaw, E., Large, D., & Burnett, G. (2020b). Driver training for future automated vehicles: Introducing CHAT (CHeck, Assess, Takeover). RAC Foundation, London, 2020.

Shepardson, D. (2020). NHTSA opens probe into fatal Tesla crash in Indiana. Automotive news. Available at: <https://www.autonews.com/regulation-safety/nhtsa-opens-probe-fatal-tesla-crash-indiana>. (Accessed: 27/07/2020).

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making?. *International Journal of Human-Computer Studies*, 51(5), 991-1006.

Society of Automotive Engineers (SAE) (2021). J3016A: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. Available at: <https://www.sae.org/blog/sae-j3016-update>

Stanton, N. A. & Marsden, P. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science*. Elsevier, 24(1):35–49.

Sullivan, J. M. et al. (2016). Literature Review of Behavioral Adaptations to Advanced Driver Assistance Systems. Available at: <https://trid.trb.org/view/1445984> (Accessed: 9 July 2019).

Taylor, R.M., (2017). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situation Awareness*, 111-128. Routledge.

Teh, E., Jamson, S., Carsten, O., & Jamson, H. (2014). Temporal fluctuations in driving demand: The effect of traffic complexity on subjective measures of workload and driving performance. *Transportation Research Part F: Traffic Psychology and Behaviour*, 22, 207-217.

Tesla Company (2019). *Model S Owner's Manual*, 208.

- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1), 97-136.
- Trösterer, S., Meschtscherjakov, A., Mirnig, A. G., Lupp, A., Gärtner, M., McGee, F., & Engel, T. (2017). What we can learn from pilots for handovers and (de) skilling in semi-autonomous driving: An interview study. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 173-182).
- Victor, T. W. et al. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors*, 60(8):1095–1116.
- Walker, G. H., Stanton, N. A. & Salmon, P. (2016). Trust in vehicle technology. *International Journal of Vehicle Design*, 70(2):157.
- Wang, Z., Bergin, C. & Bergin, D. A. (2014). Measuring engagement in fourth to twelfth grade classrooms: The classroom engagement inventory. *School Psychology Quarterly*. American Psychological Association Inc., 29(4):517–535.
- White, H., Large, D. R., Salanitri, D., Burnett, G., Lawson, A., Box, E. (2019). Rebuilding Drivers' Situation Awareness During Takeover Requests in Level 3 Automated Cars. *Ergonomics & Human Factors*, Stratford-upon-Avon, UK.
- Wickens, C. D., Lee, J.D., Liu, Y., Gorden-Becker, S. (2014). *An Introduction to Human Factors Engineering* (2nd Edition). Essex: Pearson Education Limited.
- Young, M. S. & Stanton, N. A. (2007). What's skill got to do with it? Vehicle automation and driver mental workload. *Ergonomics*. Taylor & Francis, 50(8):1324–1339.
- Zeeb, K., Buchner, A. & Schrauf, M. (2016). Is takeover time all that matters? The impact of visual-cognitive load on driver takeover quality after conditionally automated driving. *Accident Analysis and Prevention*. Elsevier Ltd, 92:230–239.

Appendices

Appendix A Training Evaluation Questionnaire (TEQ)

Instruction: Please tick the box that best represents how much you agree with each of the statements below, where 1= strongly disagree and 5 = strongly agree

Statement	1 (Strongly disagree)	2	3	4	5 (strongly agree)
I felt interested					
I felt bored					
I paid attention to the things I needed to remember					
I formed new questions in my mind as I participated in the training					
I did not want to stop at the end of the training					
I asked myself questions as I went along to make sure the training made sense to me					
I was 'zoned out' and not really thinking during the training					
I let my mind wander during the training					
The skills I learned during the training will be helpful in knowing what to do in real-life					
I would want to undertake this type of training again					
I better understand the importance of knowledge about automated vehicles after completing this training					
I take training requirements for automated vehicles more seriously after completing this training					

(Adapted from Wang, Bergin and Bergin, 2014)

Appendix B Extract from Behavioural Training – the CHAT standard operating procedure

“CHAT stands for **Check 360, Assess, Takeover**.

Check 360 represents the checks you should make all around you both inside and outside the vehicle to make sure you are ready to drive.

Check yourself – are you ready to drive? Are you feeling tired? Your first priority action, must be to put non-driving tasks, such as your phone or tablet down, out of the way and in a safe place where they cannot distract you. You may need to re-set your seat if you have moved it during automated driving mode, check you can safely reach the pedals, and switch off any internal lights. Make sure you are comfortable to drive. Once you are in the right position for driving you need to turn your attention to the road. You will need check all around you,

Check for hazards,

Check all the mirrors, and

Check your blind spot. Look to see where the vehicles and other road users are all around you. How busy is the road, are there vehicles in front of you, behind you, to the right or left of you, in your blind spot? It is important for you to know what else is on the road with you before you take over control of the vehicle movements.

Assess represents the essential assessment you need to carry out to be able to make decisions to maintain safe driving movements once you have taken over control of the full driving task.

You must **assess** your position,

the road,

the situation,

and the next steps you might need to take, both in the immediate situation and the journey.

Look to see where you are in comparison to the vehicles around you. Assess how fast the vehicles around you are going, whether they are approaching or moving away from you. Where are you on the road? Is it single or multi-lane? What vehicles are you approaching and will you need to take action such as braking or changing lanes to avoid a collision? What are the road conditions like? Are there any bends in the road that you will need to navigate safely, is the speed limit about to change, is the road wet, is it foggy. These things will all affect how you drive and your decisions, for example, to drive slower than you would if the conditions were clear. What are the next steps you need to take along your route? Are you approaching a junction or a roundabout? You need to think about things like, which direction you will need to need to turn, which lane you need to be in, which exit you need to take.

Once you have checked and assessed all around you to re-engage with the monitoring aspect of the driving task you will be ready to takeover control of the vehicle's movements."

Appendix C Operational Training User Manual

Autocar3: Automated Driving User Manual

How It Works

Your Autocar3 includes the following automated driving components that actively monitor the surrounding roadway:

1. A camera is mounted above the rear license plate.
2. Ultrasonic sensors are located in the front and rear bumpers.
3. A camera is mounted in each door pillar.
4. Three cameras are mounted to the windshield above the rear view mirror.
5. A camera is mounted to each front fender.
6. Radar is mounted behind the front bumper on the side of the vehicle.

Autocar3 is also equipped with high precision electronically-assisted braking and steering systems.

Features

Automated driving operates by simultaneously engaging the following features. It is designed to reduce driver workload and allow the use of non-driving activities:

- Autosteer
- Advanced cruise control
- Advanced roadway monitoring

Automated driving mode is a hands-free feature. When engaged, automated driving mode does not require you to hold the steering wheel, allowing you to engage in tasks and activities not related to driving, such as the use of mobile phones. However, when manual driving mode is engaged you must hold the steering wheel at all times and all traffic legislation regarding the use of non-driving activities applies as usual to manual vehicles.

As automated driving is not available in all circumstances, Autocar3 requires that you are ready to respond to a takeover request from the vehicle and remain prepared to resume full manual control of the vehicle at any time.

Autosteer

Automated driving mode intelligently centres Autocar3 in its driving lane when cruising at a set speed. Using the vehicle's camera(s), the radar sensor, and the ultrasonic sensors, automated driving mode detects lane markings

and the presence of vehicles and objects for assisting you in steering Autocar3.



Warning: Autosteer is intended for use only on dual carriageways and motorways with a fully attentive driver. When using Autosteer, be mindful of road conditions. Do not use Autosteer on residential streets, in construction zones, or in areas where bicyclists or pedestrians may be present. Always be prepared to respond to a takeover request from the vehicle. It is the driver's responsibility to be in control of Autocar3 at all times. Failure to follow these instructions could cause damage, serious injury or death.

Advanced Cruise Control

With Autocar3's automated driving mode, the forward-looking cameras and the radar sensor are designed to determine when there is a vehicle in front of you in the same lane. If the area in front of Autocar3 is clear, automated driving maintains a set driving speed. When a vehicle is detected, automated driving mode is designed to slow down Autocar3 as needed to maintain a selected time-based distance from the vehicle in front, up to the set speed. When cruising behind a detected vehicle, Advanced Cruise Control accelerates and decelerates Autocar3 as needed to maintain a following distance, up to the set speed.

Automated driving also adjusts the cruising speed when entering and exiting curves.

Note: When automated driving is actively slowing down Autocar3 to maintain the selected distance from the vehicle ahead, brake lights turn on to alert other road users that you are slowing down.



Warning: Automated driving may occasionally cause Autocar3 to brake when not required or when you are not expecting it. This can be caused by closely following a vehicle ahead, detecting vehicles or objects in adjacent lanes (especially on curves), etc.



Warning: Automated driving mode cannot operate in all circumstances. Always be ready to respond to a takeover request from you vehicle and be prepared to take over full manual control. Depending on automated driving in all circumstances can result in serious injury or death.

Advanced Roadway Monitoring

Automated driving has the capability to monitor the roadway and automatically apply the brakes when needed. Automated driving is primarily intended for driving on dry, straight roads, such as dual carriageways and motorways. It should not be used on town or city streets. Automated driving is designed for your driving comfort and convenience, it is not able to operate on all roadways and in all circumstances. It is your responsibility to stay alert, drive safely, and be ready to respond to a takeover request from your vehicle

at all times. Always be prepared to take full manual control of the vehicle, in the event that the vehicle gives notice that automated driving needs to disable and you are required to resume manual control of the vehicle. Failure to do so can result in serious injury or death.



Warning: Automated driving cannot be used on residential streets or on roads where traffic conditions are constantly changing.



Warning: Automated driving cannot be used on winding roads with sharp curves, on icy or slippery road surfaces, or when weather conditions (such as heavy rain, snow, fog, etc.) make it inappropriate to drive at a consistent speed.



When automated driving is engaged and the vehicle approaches a condition where automated driving will no longer be available for use, the vehicle will issue advanced notice for you to resume full, manual control of the vehicle.

Limitations

Many factors can impact the availability of the Automated driving mode within Autocar3, causing this mode to be disabled in certain conditions.

These include (but are not limited to):

- Poor visibility (due to heavy rain, snow, fog, sunlight etc.).
- Damage or obstructions caused by mud, ice, snow, etc.
- Interference or obstruction by object(s), mounted onto the vehicle (such as a bike rack).
- Obstruction caused by applying excessive paint or adhesive products (such as wraps, stickers, rubber coating, etc.) onto the vehicle.
- Narrow or winding roads.
- A damaged or misaligned bumper.
- Interference from other equipment that generates ultrasonic waves.
- Extremely hot or cold temperatures.



Warning: The list above does not represent an exhaustive list of situations where automated driving is not available for use. Never depend on these components to keep you safe. It is the driver's responsibility to stay alert, drive safely, and be ready to take full control of the vehicle at all times.

Engaging Automated Driving mode



You must ensure you are in the correct lane for automated driving before activating automated driving mode.

To use Automated driving, you must be driving at least 30 mph.

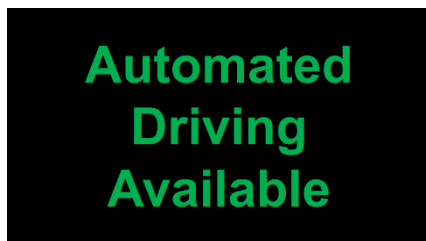
All alerts and information in relation to the driving modes and takeover requests will be displayed on the in-vehicle display screen, which is located in the area on the diagram below:



When the vehicle is in manual driving mode, the in-vehicle display screen will show the following visual alert:



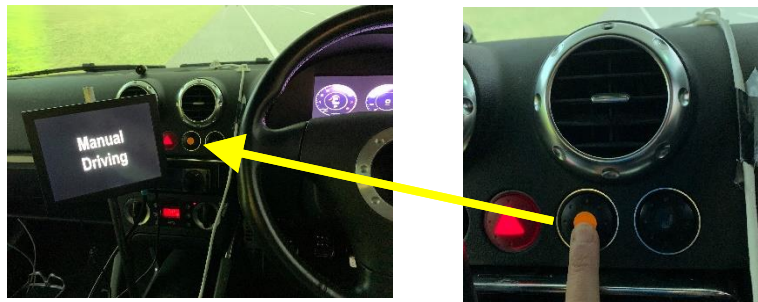
When automated driving is available the in-vehicle display screen will show the following visual alert with accompanying 'AUTOMATED DRIVING



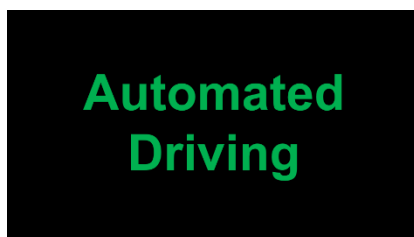
AVAILABLE' audio alert.

Note: when this alert is shown on the in-vehicle display screen, automated driving is available but is not actively controlling your vehicle movement until you activate it.

To activate automated driving mode push the orange button located above the radio as shown below:



When automated mode is engaged, the in-vehicle display will show the following visual alert with accompanying 'AUTOMATED DRIVING ENGAGED' audio alert.



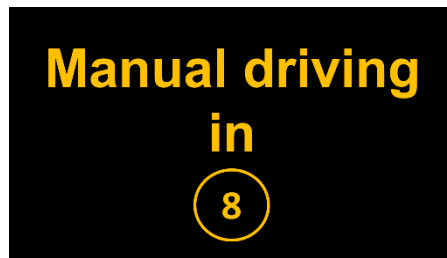
Takeover Requests and Resuming Manual Driving mode

If Autocar3 approaches a situation where automated driving will no longer be available, the vehicle will request that you takeover full manual control of the vehicle via audio-visual alerts on the in-vehicle display screen. You will receive a takeover request 60 seconds ahead of automated driving being disengaged. This time is for you to prepare to resume full manual driving mode of the vehicle. The following series of alerts will proceed when Autocar3 issues a takeover request:

When the takeover request is initiated, the in-vehicle display will show the following visual alert with accompanying 'PREPARE TO DRIVE' audio alert. This alert will be displayed for 50 seconds, an orange timer bar is displayed towards the base of the screen to give a visual indication of the countdown. 50 seconds is reached once the bar reaches the other side of the screen:



When 50 seconds has passed, the in-vehicle display will show the following visual alert and will show number counting down from 10 to 1. An audio alert will sound as the display appears, then again at numbers 3, 2 and 1.



The alert will be displayed for 10 seconds. At the end of these 10 seconds, automated driving mode will dis-engage and the display will switch to read Manual Driving mode to signal the point of manual mode engagement. Manual mode engagement will be accompanied by a longer audio alert sound and the visual display 'Manual driving'.



Note: When automated driving is dis-engaged, regenerative braking slows down Autocar3 in the same way as when you move your foot off the accelerator when driving without Advanced Cruise Control, until it becomes stationary.



Warning: Automated driving mode may issue a takeover request, in the following situations:

- Weather changes that are predicted to cause poor visibility (due to heavy rain, snow, fog, etc.).
- Daylight changes that are predicted to cause poor visibility.
- Approaching narrow, winding or residential roads.
- Approaching minimum sensor thresholds for extremely hot or cold temperatures.



Warning: The list above does not represent an exhaustive list of situations that may interfere with proper operation of automated driving mode.

When automated driving is unavailable or cancels, Autocar3 no longer centres the vehicle, drives consistently at a set speed, maintains a specified distance from the vehicle ahead or monitors the roadway. It is the driver's responsibility to be in control of Autocar3 at all times.

Use of non-driving tasks

Automated driving mode is a hands-free feature. When engaged, automated driving mode does not require you to hold the steering wheel, allowing you to engage in tasks and activities not related to driving, such as the use of mobile phones. However, when manual driving mode is engaged you must hold the steering wheel at all times and all traffic legislation regarding the use of non-driving activities applies as usual to manual vehicles.

The use of non-driving activities is permitted during automated driving mode, but normal traffic legislation for manual vehicles applies when the vehicle is in manual driving mode. It is your responsibility to stay alert, drive safely, and be ready to respond to a takeover request from your vehicle at all times.

Appendix D Critical Decision Method (CDM) 'Deepening' probe questions

Category	Questions
Cues	• What were you seeing, hearing, noticing etc.?
Information	• What information did you use in making this decision or judgement? • How and where did you get this information? • What did you do with this information?
Standard Operating procedures	• Does this fit a standard or typical driving scenario? • Is it a type of event you have training to deal with?
Goals and priorities	• What were your specific goals and objectives at this time? • What was most important to accomplish at this point in the drive?
Options	• What other courses of action did you consider or were available to you? • How was this option chosen or others rejected? • Was there a rule that you were following in choosing this option?
Experience	• What knowledge or experience was necessary or helpful in making this decision?
Mental models	• Did you create some sort of picture in your head? • Did you imagine the events and how they would unfold? • Did you imagine the possible consequences of this action?
Decision making	• What let you know that this was the right thing to do at this point? • How much time pressure was involved in making this decision? • How long did it take to actually make this decision?
What-ifs	• If a novice had been in charge at this point, what type of error might they have made and why? Would they have noticed what you noticed? Would they have known to do x? • If [key feature] had been different, what impact would it have had on your decision/assessment/actions/plan? • What training might have offered an advantage in this situation • What knowledge, information, or tools/technologies could have helped?

(Adapted from Crandall et al., 2006)

Appendix E HMI audio-visual displays for the transition from 'Manual Driving' to 'Automated Driving' mode

	Visual display for Manual Driving mode system status of operation. Central logo was represented manual driving mode throughout all HMI displays.
	Visual display that the ODD for the ADS has been entered and Automated Driving mode is available. The change in the visual display was accompanied by an audio voice message stating, "<i>Automated Driving Available</i>". The central logo represented automated driving mode throughout all HMI displays.
	Visual display of the 3 second count down to inform drivers that Automated Driving Mode is engaging. The numbers in the centre oval changed from 3 to 1 on each second count. Each number were accompanied by an alerting tone that aligned with the countdown timing.
	Visual display for Automated Driving mode, which appeared after the 3 sec countdown. This visual display remained on the HMI until a system change (e.g. TOR) to indicate the Automated Driving mode system state. The final display was accompanied by a single voice message stating, "<i>Automation Driving Engaged</i>".

Appendix F Description and units of key data metrics recorded by the AVSimulation SCANeR software.

Data and units	Description/Definition
Longitudinal Speed (kph)	Vehicle forward travel speed in kph. In the narrative of results or discussion relating to driver errors or hazardous control responses, such as speed violations, the figurative equivalent in mph will be reported.
Lateral Speed (m/s)	Speed of vehicle lateral movement e.g. during lane changes
Longitudinal acceleration/ Lateral Acceleration (m/s ²)	Acceleration of vehicle relating to forward movement/lateral movement e.g. during lane changes
Longitudinal road position (m)	Recorded as the position of the right rear axle position of the vehicle on the simulated road surface. Accurate vehicular dimensions were used to calculate front and rear bumper positions.
Lane Gap (m)	Lateral distance of the vehicle centre to the left or right of central position of the lane (0).
Road Gap (m)	Lateral distance of the vehicle centre from the motorway central reservation. Used to identify a vehicle's position in relation to lane boundaries.
Vehicle brake (N)	Provided onset and offset data for the application of vehicle brakes (used to calculate when and for how long vehicle brakes were applied)
Steering Wheel Angle (degrees)	The angular turn (in degrees) of the steering wheel to the left or right of its central position (0).
Vehicle Automation	This provided timestamp data for when automation was engaged and disengaged.
Vehicle Indication	This provided onset and offset data for when a vehicle was indicating.
Lane ID	Used to cross check lane boundary information with Road Gap position data.

Appendix G Tables reporting Mean and Standard Deviations for the Subjective Results reported in the CHAT HMI Driving Simulator Study

Figure 8-1 Mean (M) and Standard Deviation (SD) for SART dimensions for each HMI group

	Drive 1 (medium demand)		Drive 2 (low demand)		Drive 3 (high demand)	
SART Dimensions	CHAT HMI M (SD)	Basic HMI M (SD)	CHAT HMI M (SD)	Basic HMI M (SD)	CHAT HMI M (SD)	Basic HMI M (SD)
Attentional Demand	2.5 (0.86)	2.56 (0.87)	2.11 (0.99)	2.69 (1.37)	4.06 (1.63)	4.22 (1.55)
Attentional Supply	5.08 (0.54)	4.46 (0.77)	4.58 (0.82)	4.13 (0.41)	4.71 (0.86)	4.00 (0.69)
Understanding	5.08 (1.01)	4.44 (1.02)	5.31 (0.87)	4.72 (0.98)	5.5 (0.59)	4.83 (0.99)
SART	4.31 (0.47)	3.88 (0.6)	4.06 (0.64)	3.88 (0.38)	4.77 (0.78)	4.25 (0.94)

Figure 8-2 Mean (M) and Standard Deviation (SD) for NASA-TLX Workload dimensions for each HMI group

	Drive 1 (medium demand)		Drive 2 (low demand)		Drive 3 (high demand)	
Workload Dimensions	CHAT HMI M (SD)	Basic HMI M (SD)	CHAT HMI M (SD)	Basic HMI M (SD)	CHAT HMI M (SD)	Basic HMI M (SD)
Mental Demand	4 (1.21)	3.67 (1.78)	3.08 (1.5)	3.25 (1.66)	4.25 (1.42)	3.58 (1.78)
Physical	2.33 (0.89)	2 (1.35)	1.83 (0.58)	2.00 (1.28)	2.58 (1.08)	2.33 (1.49)
Temporal Demand	1.75 (2.75)	2.75 (1.55)	1.75 (0.87)	2.67 (1.30)	2.5 (1.50)	3.67 (1.61)
Effort	3.33 (1.7)	2.92 (1.56)	2.42 (0.9)	2.42 (1.50)	3.5 (1.45)	3.5 (1.68)
Performance	5.42 (1.31)	4.75 (1.29)	6 (0.74)	5.17 (0.84)	5.08 (0.79)	4.5 (1.68)
Frustration	1.92 (0.9)	3 (1.48)	1.83 (0.72)	2.58 (1.08)	3.08 (1.56)	3.17 (1.59)
Total Workload	2.65 (0.72)	2.93 (0.81)	2.13 (0.61)	2.58 (0.94)	3.18 (1.05)	3.26 (1.27)

Figure 8-3 Mean (M) and Standard Deviation (SD) for TTAQ subscales for each HMI group

	Pre Drive		(medium demand)		(low demand)		(high demand)	
Trust subscales	CHAT M (SD)	Basic M (SD)	CHAT M (SD)	Basic M (SD)	CHAT M (SD)	Basic M (SD)	CHAT M (SD)	Basic M (SD)
Intent to Use	3.98 (0.32)	4.1 (0.54)	4.2 (0.36)	4.21 (0.51)	4.37 (0.37)	4.30 (0.60)	3.75 (0.37)	3.92 (0.41)
Discharge to Driver	2.8 (0.5)	3.4 (0.6)	3.5 (0.8)	3.6 (0.8)	3.7 (0.9)	3.6 (0.7)	3.7 (0.8)	3.8 (0.7)
Trust	3.3 (0.6)	3.0 (0.6)	3.8 (0.5)	3.4 (0.6)	4.0 (3.6)	3.6 (0.6)	4.0 (3.6)	3.6 (0.6)
Total trust in automation	3.44 (0.36)	3.35 (0.39)	3.66 (0.29)	3.57 (0.42)	3.77 (0.39)	3.72 (0.45)	3.63 (0.36)	3.62 (0.39)

Appendix H Tables showing Effect Sizes for Statistically Significant results from Independent T tests conducted for CHAT HMI Driving Simulator Study

Figure 8-4 Cohen's d and effect sizes for statistically significant results of SART subgroups

Drive	SART subgroup	Cohen's d	Effect size
D1	Attentional Supply	0.69	Medium
D1	Total SART	0.54	Medium
D3	Attentional Supply	0.78	Medium/Large
D3	Understanding	0.82	Large

Figure 8-5 Cohen's d and effect sizes for statistically significant results from the NASA-TLX Workload index.

NASA-TLX dimensions	Drive	Cohen's d	Effect size
Temporal Demand	D1	0.78	Large
	D2	0.82	Large
	D3	0.91	Large
Frustration	D1	0.89	Large
	D2	0.83	Large
Performance	D2	-1.06	Large

Figure 8-6 Cohen's d and effect sizes for statistically significant results of mirror glance measures before and during the lane change in Drive 3.

Glance measure	Cohen's d	Effect size
Freq T10 mirror glances (glance per min)	-1.025	Large
Total mirror glances during LC (count)	-.897	Large
Total mirror glances T10 to vehicle lane boundary cross (count)	-1.378	Large
Total mirror glance between T10 and LC offset (count)	-1.237	Large
Glance duration T10	-.721	Medium/Large
Glance duration full LC	-.544	Medium
Glance duration T10 to boundary cross (BC)	-1.046	Large
Glance duration T10 to end lane change	-1.727	Large

Figure 8-7 Summary descriptive statistics of key driver performance metrics during the drive 3 lane change for 5 Basic condition drivers categorised as 'loss of control = yes'

ID	Lane change duration (s)	Mean speed (kph)	Max speed (kph)	Mean Lat speed (m/s)	Max Lat speed (m/s)	Mean decel (m/s ²)	Max decel (m/s ²)	Mean lat acc (m/s ²)	Max lat acc (m/s ²)	Mean SWA (deg)	Max SWA (deg)	Abrupt onset brakes (>10m/s ³)	Comments
P1B	6.35 (12.33)	87.35	100.12	0.62 (0.39)	1.37 (0.74)	3.41	3.99	0.03 (0.03)	0.13	2.86	7.35	2	Excessive speed (>100mph) - hard brakes to pull in front of the white truck and exit the motorway.
		1.37 (0.74)				0.03 (0.03)		0.13					
P4B	2.25 (10.97)	54.07	68.33	0.18 (0.27)	0.14 (1.44)	5.93	7.40	0.15 (0.14)	0.52	50.19	90.36	5	Evasive manoeuvre (steering and brakes) uncontrolled Lane change behind L2 - LBDS/inattentional blindness
		1.44 (1.44)				0.15 (0.14)		0.52					
P6B	10.05 (8.89)	58.63	73.79	0.02 (0.16)	0.63 (1.84)	8.63	10.33	0.08 (0.19)	1.21	3.29	9.94	4	Evasive manoeuvre - pulls in towards L2, then brakes.
		1.84 (1.84)				0.08 (0.19)		1.21					
P8B	10.10 (2.16)	66.84	69.48	0.06 (0.01)	0.07 n/a	n/a	n/a	0.01 (0.01)	0.04	4.80	26.20	0	Evasive manoeuvre (steering) - Long glances away from the front of the road - no SA of L1 movement. Startled by L1 when eventually looked forwards and oversteered onto hard shoulder (lane exceedance).
		n/a				0.01 (0.01)		4.80					
P9B	12.50 (5.13)	65.04	78.66	0.04 (0.12)	0.40 (2.36)	5.87	8.04	0.03 (0.06)	0.30	6.97	23.76	4	Evasive manoeuvre (steering and brakes) - drives towards Lead 2, swerves away then makes lane change a second time.
		2.36 (2.36)				0.03 (0.06)		0.30					

Appendix J Glance duration towards Roadway - descriptive statistics

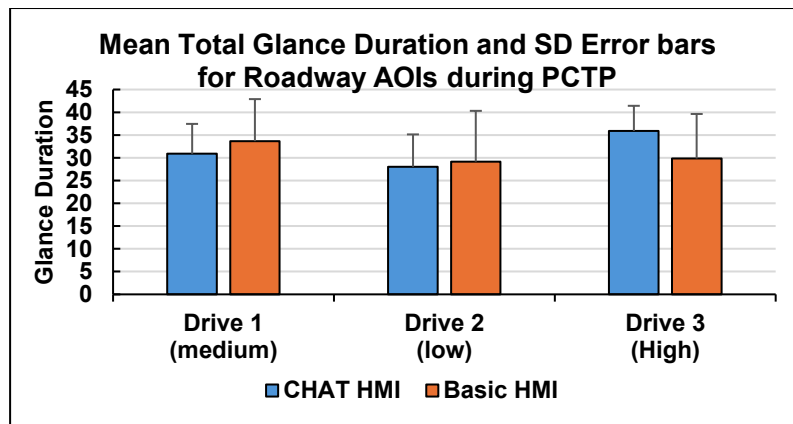


Figure 8-8 Mean glance duration and standard deviation error bars towards the roadway during the PCTP for drives 1, 2 and 3.

Appendix K PCTP T-period analysis for 'Roadway' and a comparison between T1 and T6 for AOI in glance category 'internal'

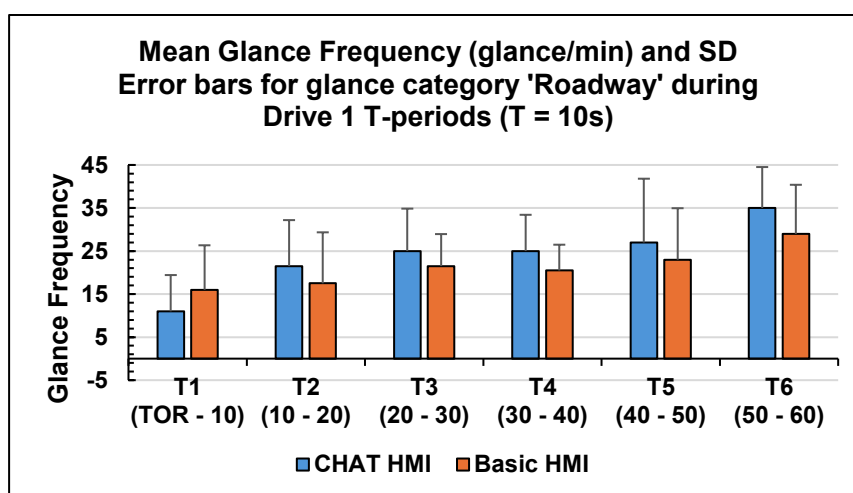


Figure 8-9 Drive 1 mean glance frequency (glance/min) and standard deviation error bars for 'Roadway' AOIs during each T period in the PCTP.

Table 8-1 Statistically significant pairwise comparisons across T-periods (T = 10s) for medium task demand total glance duration to 'Roadway' ($\alpha = 0.05$).

Paired T-periods (T = 10sec)		
Pair 1	Pair 2	p value
T1	T2	.003
	T3	<.001
	T4	<.001
	T5	<.001
	T6	<.001

Table 8-2 Pairwise comparisons for Glance Frequency and Duration to 'Roadway' for high task demand PCTP T-periods (T = 10s)

Glance Measure	Paired T-periods (T = 10sec)		p value
	Pair 1	Pair 2	
Glance Frequency	T1	T2	<.001
		T3	.009
		T4	.001
		T5	.006
		T6	.015
Glance Duration	T1	T2	.001
		T3	.010
		T4	<.001
		T5	.013
		T6	.007

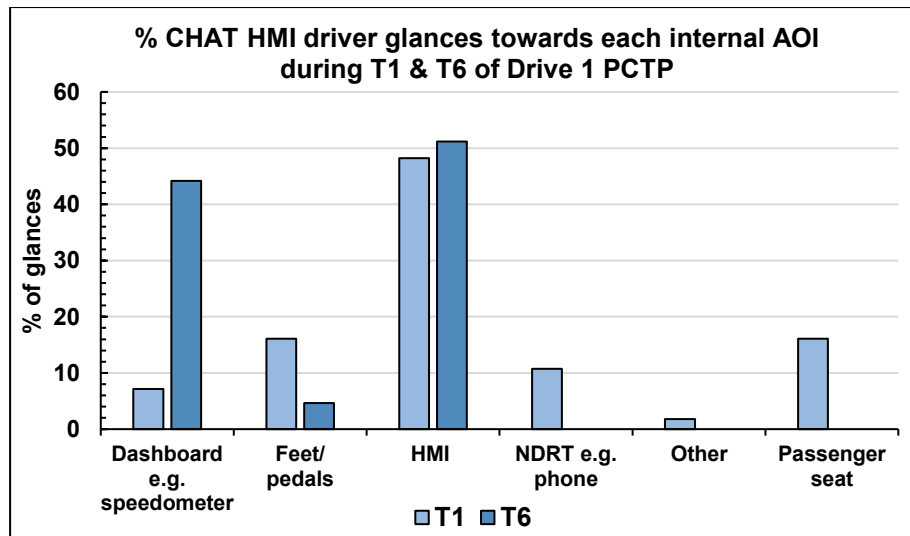


Figure 8-10 CHAT HMI driver glances % AOI within internal glance category for periods T1 (TOR – 10s) and T6 (50s to MDM).

Appendix L Descriptive Statistics and Mann-Whitney Y test scores for Drive 1 lane change

Figure 8-11 Descriptive statistics and Mann-Whitney U test results for Drive 1, lane change performance metrics

Performance Measure	CHAT HMI Md (IQR)	Basic Md (IQR)	Z	p value
Lane Change Duration	9.35 (2.488)	10.85 (6.363)	-.808	.133
Mean lateral acceleration	.0060 (.001)	.0068 (0.03)	-2.252	.024*
Max lateral acceleration	0.028 (.019)	.044 (.036)	-1.905	.057
SD lateral acceleration	.0078 (.002)	.0089 (.005)	-2.078	.038*
SD lateral speed	.0097 (.012)	.0216 (.024)	-1.963	.050*
Mean SWA	1.39 (1.185)	1.91 (2.35)	-1.386	.166
Max SWA	4.11 (6.04)	4.94 (8.932)	-1.039	.299
Min distance headway**	20.17 (21.25)	17.19 (12.89)	-0.808	.419
Min time Headway**	1.12 (1.18)	0.95 (0.71)	-.808	.419

* statistically significant ($\alpha = .05$) ** minimum headway calculated to the slow moving vehicle at the point the left edge of the ego vehicle reached the lane boundary (along the x plane).

Appendix M Table showing Mean and Standard Deviations for mirror glances at time blocks during the Drive 3 lane change

Figure 8-12 Descriptive Statistics and test scores for mirror glances 10 seconds before and during the lane change in drive 3.

Time block	Glance Measure	CHAT HMI Mean (SD)	Basic HMI Mean (SD)	t	df	P value
T10 (10s before LC onset)	Glance frequency (glance/min)	29.50 (9.386)	18.00 (12.792)	-2.511	22	.010
	Glance duration	2.76 (1.427)	1.79 (1.268)	-1.765	22	.046
Total Lane Change	Glance count	4.00 (1.651)	2.58 (1.505)	- 2.196	22	.019
	Glance duration	2.49 (1.111)	1.83 (1.287)	-1.333	22	.098
T10 to vehicle lane boundary cross	Glance count	8.00 (2.663)	4.67 (2.146)	-3.376	22	.001
	Glance duration	4.67 (1.847)	2.92 (1.481)	-2.562	22	.009
T10 and to LC offset (full LC).	Glance count	8.92 (2.539)	5.58 (2.843)	-3.029	22	.003
	Glance duration	5.25 (1.878)	3.62 (2.002)	-2.053	22	.026