The Influence of Vehicle Automation on Visual Attention and Situation Awareness

Stuart Hyde

Department of Psychology

Royal Holloway, University of London

Submitted for the Degree of Doctor of Philosophy, September 2017

Declaration

I, Stuart Hyde, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed:______________________

Date:________________________

Abstract

The increasing prevalence of vehicle automation and, in the future, fully autonomous vehicles, creates a need for research into the likely effects of this significant change to the way we drive. This thesis examines the likely impacts of increasing automation on driver attention and situation awareness. Chapter 1 provides a general overview of the topic and the research questions. Chapter 2 presents the first experiment, which measures hazard detection performance under different levels of load in a secondary task. Chapter 3 presents a second similar experiment using an older (70+ years) population. Chapter 4 presents a third experiment which again repeated the first study but this time in a young, inexperienced driving population (18-20 years). Chapter 5 then presents a direct comparison of the results of the three previous studies. Chapter 6 describes a large simulator-based study to examine the visual attention and situation awareness of drivers after they take over driving following a period of autonomous driving. Driver performance is assessed using a range of eye tracking and behavioural measures. Chapter 7 is the final empirical chapter, presenting the results of an online survey aimed at understanding people's trust and acceptance of vehicle automation. Chapter 8 constitutes a general discussion in which all study results are reviewed and their implications discussed. Overall, this thesis uses a range of different experimental methods to demonstrate that drivers' attention and situation awareness is likely to be significantly affected by differing types and levels of vehicle automation. This will be an important consideration for the designers of automated systems and vehicles in the future.

Acknowledgements

This PhD journey has been a roller coaster of emotional highs and lows, of new experiences and friendships. Consequently, there are quite a few people to thank. First and foremost, the biggest thanks goes to Polly Dalton. Your patience, humility, generous encouragement, and unbridled enthusiasm for this PhD project have undoubtedly helped drive me to the finish line. You have been the best supervisor I could wish for. Under your guidance I have learnt a great deal both professionally and academically, and I look forward to taking these skills with me in my future career and life. I wish you all the best for your sabbatical year and future research.

Alan Stevens, from TRL, stepping into an ongoing project can't be easy but you transitioned in wonderfully. Thank you for your thoughtful guidance and your kind support. You helped push me that little bit further in my critical thinking, as well as challenging me to be more precise with my communication of ideas. Andrew Parkes, thank you for your initial supervision of my PhD, which was always calm, considered and insightful, helping to set me up for my PhD journey. Simon Wilson, thank you immensely for the time and effort you put into helping me design the simulator study, and taking the time to talk through all the practical aspects of my study design. Also, thank you Lena Weaver for helping with the recruitment of the older drivers and the drivers for the simulator study. Without your time and assistance I'd probably still be trying to recruit participants!

From Royal Holloway, a big thank you to Sandra Murphy for being a friend, colleague and all round fantastic person to talk to about life, food, academia, psychology and programming. Your help in the early days of my PhD with E-prime was invaluable, and in the last few months your guidance and help with R and eye-tracking analysis probably saved Chapter 6 of my PhD from ruin! For all these things, I am truly grateful to you. To Craig Arnold, thank you for all your data wrangling help. The data from Chapter 6 was only useable after your considerable assistance getting it into shape and in a useable format. Without your expert MatLab skills I'd still be a long way from finishing. Best of luck with your PhD; you have a great team around you. Thank you to my advisor Rob Hughes, and Johannes

Zanker, Courtenay Norbury and Cat Sebastian for making my annual reviews and upgrade such invaluable experiences. Your questions, insights and enthusiasm for my research helped push me forward every year.

My friends, and, importantly family you have provided a constant backdrop of love and support that has helped me through the challenging times and revel in the good times. A special thank you to my wonderful wife, Tabitha; your love, patience and support when work was getting me down, and, your happiness (and relief!) when things were going well have helped me grow, learn and become a better person. It's hard to believe we married, bought a house, and achieved so much else in between, over the last 3 years. The future is bright, the future is ours.

Finally, to all the wonderful music, story books, podcasts (No such thing as a fish, in particular!) that motivated, provided escape and helped me retain sanity over the last three years, thank you.

This thesis is dedicated to my father David Hyde, thank you for all your love and guidance. I'm sorry you couldn't be here to see me complete my PhD, I know you would have loved discussing my work and I would have loved discussing it with you.

Publications and Dissemination

Chapter 2 was submitted for publication at Psychological Science and QJEP.

Chapter 6 is in preparation for publication.

Chapter 7 is published as a report through Transport Research Laboratory (TRL). *Attitudes to autonomous vehicles (2017)* ISBN: 978-1-910377-91-8. https://trl.co.uk/reports/attitudes-autonomous-vehicles.

Table of Contents

List of Figures

List of Tables

Chapter 1 - General Introduction

Background

Through movies, literature and sci-fi the idea of a self-driving car has been dreamt of and imagined. From Blade Runner in 1982 through to the Fifth Element in the late 90’s the fantasy and allure of the self-driving car has grown with each passing year. Yet it is only in the last decade that the reality of the self-driving car has come closer to fruition. The invention of adaptive cruise control and active steering systems in the 90’s led the way for the early acceptance of automated systems in vehicles. Latterly, in the last decade there has been a rapid race by the automotive industry to introduce new partially automated vehicle systems, such as vehicles that can park themselves, vehicles that speed up and slow down automatically in stop-and-go traffic and vehicles that can change lanes on motorways. These partially automated vehicle systems are paving the way forward for the mass production and introduction of fully automated cars on the roads of the world in the not too distant future.

The introduction of vehicle automation could be as transformational to our daily living as the internet was. Driverless and semi-autonomous vehicles could change the landscape of cities, drive revenue and jobs in British industry, and improve the safety of road travel for both drivers and pedestrians (CCAV, 2016). The average UK driver spends 235 hours per year driving and drives just over 8500 miles per year. The UK government reports that in 2014 there was on average one fatality every 178 million miles driven on UK roads. Driverless and semi-autonomous vehicles are predicted to have an important effect on the reduction of road traffic collisions and fatalities. Government and independent research has shown that more than 90% of road traffic collisions have human error as a contributory factor, implying that driverless vehicles could help to reduce death and injuries on our roads. Specifically, collision avoidance technologies such as Electronic Stability Control (ESC) and Autonomous Emergency Braking Systems (AEBS) have shown a more than 20% benefit in collision reduction. Observations such as this have led to claims that the adoption of semi-autonomous and driverless technology could save over 2,500 lives and prevent more than 25,000 serious accidents in the UK by 2030. This increased safety may reduce many common accident risks and therefore crash costs and reduce insurance premiums (CCAV, 2016; KPMG, 2015)

As driverless technology becomes more widespread, information concerning the real-world performance of driverless vehicles becomes more available. For example, Google reports that collectively their driverless cars have driven more than 500,000 miles without crashing, and in total over 1.5 million miles driverless (Waymo, 2017). The first reported causality in a semi-autonomous vehicle was recently reported by Tesla, following over 130 million miles of driving. However, this compares favourably to the fatality rate of manual driving, both in the US, where one driving-related fatality is reported on average every 94 million miles driven, and worldwide, where there is one fatality approximately every 60 million miles (Department for Transport, 2016).

The wider impact on society from driverless technology is predicted to be extensive. Driverless vehicles could enable better use of road space, for example through platooning of vehicles (vehicle groups traveling close together), the use of narrower lanes, and reduced junction stops. This would lead to improved traffic flow, reduced congestion and increased road capacity, which would in turn improve fuel economy and costs. Autonomous vehicles may also increase fuel efficiency and reduce pollution emissions, due to efficiency savings made through improved engine design and more efficient use of the road. On a more social level, driverless technology could improve the lives of disabled or older people by enhancing their mobility, giving transport access to those who currently cannot drive. This could provide independent mobility for non-drivers, and therefore reduce the need for motorists to chauffeur non-drivers, and to use subsidised public transport. The costs of driving may also be reduced, through lower insurance premiums and a more widespread use of car sharing (vehicle rental services that substitute for personal vehicle ownership). Finally, driverless technology could reduce driver stress, allowing motorists to relax, socialise and work while traveling.

The impact of driverless technology on the UK economy and UK businesses is predicted to be substantial. It is estimated that driverless technology could increase GDP by 1% by 2030, and that the economic and social benefit of connected and driverless vehicles (i.e. fewer accidents, improved productivity and increased trade) could be in the region of £51 billion per year by 2030. The UK government predicts that driverless vehicle technologies could lead to improved productivity and increased trade as UK industries capture part of a wider global market for Intelligent Mobility estimated to be worth £900bn worldwide by 2025. This influx of new business and industry for driverless vehicles could potentially create an additional 320,000 jobs in the UK by 2030, of which 25,000 would be in automotive manufacturing. Longer-term estimates predict that by 2040 the economic benefit to the UK economy could be more than double those in 2030 (£51 billion) at £121 billion (KPMG, 2015).

The spread of vehicle automation in the car industry has grown rapidly over the last 5-10 years with many newer cars having some level of automation as a standard feature, such as adaptive cruise control, collision avoidance and assisted parking. This pattern of increasing automation is likely to continue at pace over the next few decades, resulting in full automation becoming the norm well within this century. For example, both Ford and Tesla are aiming to have vehicles in production by 2020-21 in which the automated driving system monitors the driving environment with limited or no driver input (Naughton, 2017; NHTSA, 2013).

In summary, the introduction of autonomous vehicles will constitute perhaps the largest change to everyday transportation in living memory and is predicted to deliver a wide range of environmental, social and economic benefits. The technological and engineering issues of automated vehicles have been well researched, yet the psychological impacts and problems have yet to be fully addressed. As such, many important research questions remain, such as: How might overall driving performance in a semi-autonomous vehicle compare to entirely manual driving? How will visual behaviour change in comparison to driving a manual vehicle? Will automation have an impact on situation awareness and, if so, how? How will the interaction of the user and the automated vehicle differ from that of the user and a manual vehicle? What are the processes involved in resuming manual control when a highly automated vehicle needs to hand back to a human driver?

Key questions

The key questions in this thesis will centre around the effects of different levels of vehicle automation on situation awareness measured through the detection and perception of driving hazards (Chapters 2 and 5), and how these issues affect high risk driving groups, namely older drivers (Chapters 3 and 5) and novice drivers (Chapters 4 and 5). Additionally, this thesis will examine visual attention and situation awareness during the process of handover from an automated system back to a manual driver (Chapter 6). Finally, this thesis will examine some of the factors that influence the trust and acceptance of vehicle automation in the driving population (Chapter 7).

The detailed literature relating to each topic will be covered at the start of the relevant chapter. In this introductory chapter, I cover the issues that have broader relevance for the thesis as a whole, namely: levels of automation, measures of task workload, and the interaction between vehicle automation and situation awareness.

Levels of Vehicle Automation

An automated vehicle (often called a self-driving car, an automated car or an autonomous vehicle) is a robotic vehicle designed to travel between destinations with limited or no human input. The industry standard for describing the different levels of vehicle automation is the SAE J3016. The definitions are as follows:

- "Level 0 – no automation: the full-time performance by the human driver of all aspects of the dynamic driving task (includes the operational steering, braking, accelerating, monitoring the vehicle and roadway, and tactical responding to events, determining when to change lanes, turn, use signals, aspects of the driving task, but not the strategic determining destinations and waypoints aspect of the driving task), even when enhanced by warning or intervention systems
- Level 1 – driver assistance: the driving mode-specific (type of driving scenario with characteristic dynamic driving task requirements e.g.,

motorway merging, high speed cruising, low speed traffic jam, etc.), execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the human driver performs all remaining aspects of the dynamic driving task

- Level 2 – partial automation: the driving mode-specific execution by one or more driver assistance systems of both steering and acceleration/deceleration using information about the driving environment and with the expectation that the human driver performs all remaining aspects of the dynamic driving task
- Level 3 – conditional automation: the driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene
- Level 4 – high automation: the driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task, even if a human driver does not respond appropriately to a request to intervene
- Level 5 – full automation: the full-time performance by an automated driving system of all aspects of the dynamic driving task under all roadway and environmental conditions that can be managed by a human driver"

(SAE, 2014)

The experiments in this thesis relate to different levels of automation. Chapters 2-5 report results from a laboratory-based experimental paradigm involving hazard perception under differing levels of load. This paradigm relates most closely to situations of level 2 automation, in which the driver is expected to remain fully engaged with the driving environment. Chapter 6 reports a simulator experiment examining the process of handover between an autonomous system and a human driver and thus relates most closely to situations of level 3 automation.

Measures of Task Workload

An important facet of the effect of vehicle automation on behaviour is its specific effect on task workload. This will constitute a major topic of investigation in Chapters 2-5 of this thesis. Indeed, workload and situation awareness (considered in the following section) have been argued to be the two most important factors in predicting driver performance and safety (Parasuraman, Sheridan, & Wickens, 2000; Young & Stanton, 2002b). Workload can be defined as the physiological and mental demands that occur while performing a task or a combination of tasks. Hart and Staveland (1988, p.140) specifically define that "Workload is not an inherent property, but rather it emerges from the interaction between the requirements of a task, the circumstances under which it is performed, and the skills, behaviours, and perceptions of the operator".

In the human factors literature, workload is often assessed using a variety of a unidimensional instrument to measure subjective mental workload. The RSME consists of a line with a length of 150 mm marked with nine anchor points, each accompanied by a descriptive label indicating a degree of effort (e.g. 0 = absolutely no effort, 40 = some effort, 80 = great effort, 120 = extreme effort). The NASA task load index (NASA-TLX) (Hart & Staveland, 1988) is a six-dimension measure of mental workload: mental demand, physical demand, temporal demand, performance, effort, and frustration. The Subjective Workload Assessment Technique (SWAT) (Reid & Nygren, 1988) is a subjective rating technique to assess workload that uses three levels: low, medium, and high, for each of three dimensions of load time: mental effort load, and psychological stress load. Finally, the workload profile (WP) (Tsang & Velazquez, 1996) asks the participants to report the proportion of attentional resources used after they have experienced all of the tasks to be rated. These measures of workload are all subjective, relying heavily on the recall of task workload by participants, which can be influenced by cognitive biases and memory recall (for a detailed review of the use of subjective rating scales in the applied workload literature see Annett, 2002; Young, Brookhuis, Wickens, & Hancock, 2015).

In contrast, there is a large and influential body of theoretical research into selective attention that has examined task load and its effect on information processing in fine-grained detail. One of the most influential is the load theory of selective attention and cognitive control (Lavie & Tsal, 1994; Lavie, 2005; Lavie, Hirst, de Fockert, & Viding, 2004). There are two aspects to the theory, perceptual load and cognitive (executive) control load. According to perceptual load theory, perceptual processing has a system capacity limit but processing proceeds automatically and in parallel on all the information within the system's capacity until capacity runs out. According to the theory, the perceptual workload in the relevant task dictates whether or not a task-irrelevant stimulus is processed. More specifically, under low perceptual load in the relevant task, additional processing capacity 'spills over' to the processing of task-irrelevant stimuli. In contrast, under high load conditions the demands of the relevant task exceed the system capacity limits, and information processing is inexorably selective as a result. Therefore, there is little to no interference by task irrelevant stimuli under high perceptual load, as there is no spare capacity to process the irrelevant information. It is important to note that the distinction between a low and high task workload is relative, rather than absolute. For example, Rees Frith and Lavie specified that increasing perceptual load entails "increasing the number of items in a display or increasing the processing requirement for the same number of items" (Rees, Frith, & Lavie, 1997, pg. 1618). Thus, unfortunately, there is currently no way of quantifying the extent of the perceptual load imposed by a particular task in order to allow comparison with other tasks.

Load theory also specifies that executive control functions (e.g. working memory) act to maintain stimulus processing and response goals during the course of a task being performed. The process is an active means of control, which allows the maintenance of stimulus processing goals in order to minimise the distraction effect of task irrelevant stimuli even after they have been perceived (e.g. in a task involving low perceptual load, which leads to spare processing capacity, resulting in a distractor effect). Evidence for this proposal typically comes from studies using a dual task design in which participants are asked to hold information in mind whilst also completing a selective attention task. A commonly used task is the Eriksen flanker task, in which participants are tasked with identifying a single target letter whilst ignoring a concurrently presented distractor letter (Eriksen & Eriksen, 1974; Eriksen & Spencer, 1969). The distractor letter is either congruent or incongruent with the target letter. In the secondary task, participants perform either a low working memory load task (e.g. memorise one digit) or a high working memory load task (e.g. memorise six random digits). The results of this type of study typically show that distractor congruency effects are greater under conditions of high working memory load compared to low working memory load (e.g. Lavie & Dalton, 2014). According to the theory, when working memory load is high in this paradigm, reduced executive capacity is available for maintenance of current goals and task priorities, leading to increased interference effects by task irrelevant distractors (Lavie, 2005).

In conclusion, load theory predicts that increasing executive control load in a secondary task increases the effect of distractor interference in the primary task, whereas increasing perceptual load in the primary task decreases the effect of distractor interference within that task. The executive control aspect of load theory is likely to be highly relevant to the current work. However, the definition of task-relevant or -irrelevant stimuli in the context of a visual driving scene is challenging. In a typical driving scene, there are large numbers of task-relevant and -irrelevant stimuli, which are not easily distinguishable from one another, in addition to a number of secondary tasks (e.g. speed control). In most instances, a large number of elements in a typical driving scene could be a classified as potential hazards or useful information, making them potentially task-relevant depending on how the scene develops. In contrast, in a typical executive control load experiment there is usually a small number of clearly distinguishable task-relevant and -irrelevant stimuli, and a single secondary task (e.g. digit span recall). Nonetheless, this is an important endeavour, because the findings of load theory have demonstrated that differing types of load can have vastly differing effects on performance – a finding that has largely been overlooked within the human factors literature.

Automation and Situation Awareness

Situation awareness in the context of vehicle driving involves the perception of environmental elements (e.g. traffic speed), the knowledge of the elements' meaning, and the projection of their status after some variable has changed (either unpredictable, such as traffic braking, or predetermined, like traffic lights changing; Endsley, 1988). In simpler terms situation awareness is ''knowing what's going on so you can figure out what to do" (Adam, 1993, p.319). Understanding the interaction between vehicle automation and situation awareness is critically important, as there are two clear predictions that can be made. On the one hand, automation could positively influence situation awareness by freeing up cognitive resources, thereby increasing awareness and speeding response times to hazardous or potentially hazardous situations. Indeed, there is an overwhelming consensus from the cognitive psychology literature that reducing the number of concurrent tasks to be carried out should improve performance on the remaining tasks (e.g. Darling & Helton, 2014; Ettwig & Bronkhorst, 2015; Pashler, Johnston, & Ruthruff, 2001; Pashler, 1994; Watanabe & Funahashi, 2014). From this perspective, the clear prediction is that increasing vehicle automation should improve situation awareness. On the other hand, increasing automation could negatively affect situation awareness through cognitive under-load, as drivers become disengaged with the task and are unable to respond as quickly or effectively to driving hazards. Research in this area is still at a relatively early stage and at present there is no clear consensus as to whether increasing automation improves or reduces situation awareness.

On the one hand, there is evidence showing that high levels of automation can reduce situation awareness. For example, in a meta-analysis of the effects of adaptive cruise control (ACC) and highly automated driving (HAD) systems on situation awareness, de Winter, Happee, Martens and Stanton (2014), found that HAD systems increase drivers' susceptibility to drowsiness and mind wandering, such that response times to hazards are slower than during manual driving. Specifically, in HAD situations drivers were less likely to spot the sudden appearance of hazards (e.g. sheep, police cars) at the side of the road, as compared to manual driving (63% vs. 77%). Importantly, HAD systems and ACC also induced long response times and an elevated rate of near-collisions in critical events when compared to manual driving.

In line with these findings, there is some evidence of a detrimental effect of task under-load on situation awareness. Endsley and Kaber (1999) examined the effect of different levels of automation on situation awareness in a simulated system-monitoring task. Participants had to monitor information using a complex, computer-based dynamic control task, under varying levels of automation and in which several automation failures occurred. Examples of the tasks undertaken include object collision avoidance, and location and selection of objects. There were ten levels of automation, ranging from manual control to full automation. Situation awareness was measured using the Situation Awareness Global Assessment Technique (SAGAT) developed by Endsley, 1988, a measure where the driving simulation is temporarily frozen and the screen(s) are blanked. During the simulation freeze, drivers fill out a brief question sheet probing them about their awareness of the driving environment. Task workload was measured using the NASA-TLX (Hart & Staveland, 1988). The results showed that lower levels of automation benefited situation awareness performance the most, whilst under high levels of automation situation awareness performance was poorest, particularly if there was a system failure. However, the results of this study are difficult to generalise to the driving domain, as none of the tasks were driving-based. The task of driving is one that is highly dynamic, from controlling the vehicle to the constantly changing environment outside the vehicle, whereas the task of monitoring readouts and information from a dynamic display is arguably more static and predictable in nature than driving. Nonetheless, the results do demonstrate the general principle that cognitive under-load can be detrimental to performance, despite freeing up cognitive resources. Furthermore, Parasuraman and Riley (1997) also demonstrated that under-load as a result of high levels of automation can cause system users to become disengaged with a task and thus slower to respond to critical events.

In line with these findings, there is also evidence that high levels of automation can negatively impact patterns of visual attention to the driving scene. Using a driving simulator and eye tracking, Merat, Jamson, Lai, and Carsten (2012) found that levels of automation (none, ACC and HAD) and a secondary task (e.g. reading, watching a DVD, changing the radio) affected the allocation of attention to the roadway. Specifically, drivers in HAD vehicles were less likely to look at the centre of the roadway, where the important driving information is, compared to manual drivers (53% vs. 72% of the time). HAD drivers were also likely to spend a larger proportion of the time in the vehicle engaged in secondary tasks compared to manual driving (43% vs. 22% of journey time engaged in watching a DVD). This shift in attention away from the centre of the roadway under HAD conditions is likely to negatively impact on situation awareness.

Overall, the research described above suggests that increased automation can decrease drivers' situation awareness. However, there are a growing number of studies finding the opposite result: that increased automation can in fact be beneficial to situation awareness compared to manual driving. For example, in a recent simulator study that used non-driving related stimuli (pointed arrows) to assess situation awareness, drivers responded faster to the arrows during HAD than during manual driving (1800 ms vs. 1940 ms; De Winter, De Winter, Stanton, Price, & Mistry, 2016). This does suggest that vehicle automation can benefit visual performance. However, the use of non-driving related stimuli for the assessment of situation awareness makes these findings hard to generalise. For example, the non-driving related stimuli were abruptly-presented and incongruent with the scene and are thus likely to have been much more salient than driving related stimuli, which typically have a gradual onset and little saliency enhancement compared to the surround (McCarley, Steelman, & Horrey, 2014).

However, similar effects have been found using more ecologically valid tasks. For example, in a driving simulator study Ma and Kaber (2005) observed that drivers had better situation awareness for ACC driving than for manual driving. Mean SAGAT scores were 77% for ACC vs. 52% for manual driving, meaning that when probed ACC drivers had better recall of the driving situation on average compared to manual drivers. In a subsequent follow up study, Ma (2006) replicated the same pattern of results with a larger sample than in the original study. However, one of the main drawbacks of the SAGAT is that it measures people's memory for driving scenes, rather than their online awareness of the scene. Any inaccuracies in responding could therefore be explained in terms of memory failures, rather than reductions in real time situation awareness. For this reason, it is also important to develop measures of online awareness failure, and this is one of the aims of the current work (see Chapters 2-5).

This approach was taken in a real-vehicle study by Davis, Animashaun, Schoenherr and McDowell (2008). Twelve civilian drivers in semi-autonomous military vehicles were tasked with spotting targets (e.g. oil barrels, cones and rubbish bins) placed in the driving environment. The control of the vehicle was varied between semi-autonomous and manual control at different speeds. Drivers responded to the detection of a target by pressing a button on the steering wheel of the vehicle, in addition to verbally reporting the location and type of target detected. The drivers' detection of environmental objects was significantly better in the semi-automated compared to the manual driving condition (47% vs. 39% of objects). The semi-autonomous driving system was also associated with significantly better performance compared to the manual driving task in several other aspects of situation awareness, such as enhanced target identification and improved performance for unanticipated stops (e.g. emergency braking).

There are also a small number of studies advancing the idea that vehicle automation compared to manual driving has no effect on situation awareness. In a driving simulator based study that examined the effect of manual driving and two levels of automation (ACC and HAD) on drivers' responses to a critical event (a car pulling out from a parking space), no significant difference in response time to the critical event were found between manual or automated driving. Specifically, response times to the critical event were 3000ms, 2800ms and 2200ms for manual, ACC, and HAD, respectively (Martens, Wilschut, & Pauwelussen, 2008). However, the 800ms difference between the manual and the HAD condition might suggest that there could be an effect which might be masked by a lack of statistical power in this study. Nevertheless, there are several other studies reporting null effects of automation on situation awareness. For example, when lorry drivers experienced a HAD system and manual driving scenario in a driving simulator, there was no difference in response time to an unexpected deceleration of a vehicle in front without brake lights when the HAD system failed compared to manual driving (Lank, Haberstroh, & Wille, 2011). Similarly, in a driving simulator study Young (2000) found no differences between ACC and a HAD system regarding the number of drivers who responded to automation failure occurring without a warning signal at the same time as the critical situation, which was a car in front braking suddenly (27 vs. 25 out of 44 per condition responded to the failure).

In summary, it seems that the evidence for vehicle automation improving situation awareness is mixed. Some studies demonstrate a clear benefit of automation on situation awareness in the form of detecting more hazards and responding to them faster, whereas there are a number of studies finding that automation has negligible or even negative impacts on situation awareness compared to manual driving, particularly in HAD system vehicles. This mixed pattern of existing results highlights a critical need for ongoing research in this area, with the aim of clarifying the links between automation and situation awareness. This is a central aim of this thesis, starting with the new experimental paradigm that I will describe in Chapter 2.

Chapter 2 – The Effects of Workload on the Perception of Driving Hazards

Introduction

Until recently, driving a vehicle has been a complex activity, involving coordination of a range of simultaneous tasks. However, the experience of driving is set to change dramatically over the next decade, with significant increases in the prevalence of automated driving assistance systems (NHTSA, 2013). Indeed, many drivers are already using adaptive cruise control and/or active steering, reducing the need to monitor and respond to speed and lane positioning. Given the rapid pace of technological change in this area, it is likely that further substantial reductions in the demands associated with driving will arise in coming years. It is therefore essential to understand how these changes in demand will impact on those critical aspects of driving performance that are likely (at least in the near future) to be left to the driver. In Chapters 2-5, I focused on the anticipation and perception of driving hazards.

As described in Chapter 1, there is an overwhelming consensus from the cognitive psychology literature that reducing the number of concurrent tasks to be carried out should improve performance on the remaining tasks (e.g. Darling & Helton, 2014; Ettwig & Bronkhorst, 2015; Pashler, Johnston, & Ruthruff, 2001; Pashler, 1994; Watanabe & Funahashi, 2014). From this perspective, the clear prediction is that increasing vehicle automation should improve hazard perception. However, research in the driving literature does not reflect this consensus. Instead, the question of whether automation (vs. manual driving) can benefit hazard perception has produced a very mixed pattern of results (described in full in Chapter 1). It is difficult to resolve this pattern of findings with the overwhelming level of consensus that is found in the cognitive psychology literature. The purpose of the present study was to bring these literatures together, using a controlled cognitive psychology methodology to investigate the influence of task demand on hazard perception. The underlying assumption of this approach is that increasing the level of automation of the driving task will result in reduced overall task demands. Thus, this study uses a secondary task as a workload demand manipulation as a proxy for variations in automation levels.

Observers viewed videos of driving scenes and reported the location and nature of a potential hazard in each scene, while also completing a concurrent number probe task, which varied in its demand. There were three levels of demand in the probe task: a no load condition in which only the hazard perception task was completed; a low load condition in which a simple response to the number probe was required, along with the hazard response; and a high load condition in which a two-alternative forced-choice (2AFC) response was made to the number probe, along with the hazard response. The clear prediction from the existing cognitive research is that increasing the demand of the probe task should reduce performance in the hazard detection task.

As described in Chapter 1, there is a large literature examining the impact of variations in task load on performance (e.g. Lavie, 2005). This research has recently been extended to the specific context of driving. Murphy and Greene (2017) used a driving simulator to assess the effect of perceptual and cognitive load on drivers' visual search. In line with load theory, they found that high perceptual load significantly increased inattentional blindness and deafness for stimuli that were both relevant and irrelevant to driving (e.g. pedestrians or the sound of a car horn). High perceptual load also increased RTs to driving hazards (1129 ms under high load vs. 993 ms under low load) and reduced response accuracy to the hazard (97.8% under high load vs 98.6% under low load). Finally, drivers also had more collisions with parked vehicles under high load (1.36 crashes) than low load (0.14 crashes). Additionally, Murphy and Greene compared the effects of cognitive control load on driver attention in more detail, comparing verbal and visuo-spatial working memory load. Interestingly, but consistent with load theory, increasing cognitive load actually reduced levels of inattentional blindness for an unexpected object in the driving environment.

In the light of these findings, it was important to ensure that any impacts of the task demand manipulation in the current study could not be attributed to differences in perceptual load between the different task demand conditions. For this reason, I chose only hazard perception videos that contained the most salient hazards, in an attempt to control for the number of task-relevant stimuli contained in each video. Although this of course did not ensure that the perceptual demands were accurately matched across the videos, I also allocated videos into task conditions at random, with different allocations for each participant. Thus, there should have been no systematic differences between the perceptual demands of the hazards video in each condition, because these varied from participant to participant. I also controlled the secondary number probe task as far as possible, matching precisely the manual and visual demands across the high and low load conditions in order to allow a pure manipulation of executive control load. This is in contrast with the studies reported in the applied literature in which the task workload is often weakly defined and poorly controlled, potentially confounding measures with other factors, such as additional tasks.

Method

Participants

Thirty students were recruited for the study from the Royal Holloway university population (mean age = 20.7, SD = 2.3, 5 males). The sample size was derived from prior research, constituting a typical size for studies in this area (e.g. Galpin, Underwood, & Crundall, 2009). 28 participants were right-handed. All had a full driving licence (24 UK, 5 E.U and 1 Indian driving licence) and a minimum of two but no more than five years driving experience (mean experience in years = 2.7, SD =.8). The average miles driven per year was 3215 (SD = 3093). Participants were naïve to the purpose of the experiment, and none had previously taken part in driving experiments. All UK-based drivers had taken the hazard perception test as part of their driving theory test. All participants had normal or corrected-to-normal visual acuity, gave informed consent before participating, and were compensated with £10 for participation in the study. All procedures were reviewed and approved by the Departmental Ethics Committee.

Apparatus

Visual stimuli were presented on a 19 inch LCD monitor (Samsung SyncMaster 940N) with a resolution of 1280 x 1024 pixels, a refresh rate of 60 Hz and an unrestrained viewing distance of approximately 65 cm. Responses were recorded using a computer keyboard and two individual foot pedals (Treadlite II, Linemaster Switch Corp).

Stimuli

Stimulus presentation was controlled through E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Visual stimuli were supplied by a2om and are direct examples of hazard perception videos as used by the UK Driver & Vehicle Standards Agency (DVSA) in their training material. All videos contained naturally-occurring traffic situations, recorded in the first person and showing the roadway ahead but no driving mirrors. 116 videos were evaluated for suitability. Any that had more than one potential hazard and/or had a potential hazard occur in the first 10 seconds were excluded. The remaining 103 videos were then reviewed by two evaluators (one was the author, the other was not associated with the project, mean age = 29 years, mean driving experience = 8 years) who judged the first hazard onset, the nature of the hazard and the spatial location in each video. Of the remaining 77 clips for which the two evaluators agreed on all these criteria, 69 were chosen at random for use in the study. The mean video duration was 33s (S.D = 6s) and the mean hazard onset was 16s (S.D = 4s). An intraclass correlation (ICC = .85, 95% CI = .77 - .90) indicated a high level of agreement between the evaluators concerning the first hazard onset. The videos depicted four different types of driving environment (Motorway = 5, Rural = 28, Urban = 19, Suburban = 17).

The number probe in the low load condition was a white number 30 presented (0.6° x 0.6° of visual angle) on a black background and positioned in the bottom right hand corner of the screen. The number probe in the high load condition changed in identity from trial to trial but was presented with the same specifications as the number in the low load condition.

Design and Procedure

The main task for participants was to detect and respond to the presentation of the driving hazard that occurred in each video. A driving hazard was defined as any event that may result in the driver having to take some action, such as changing speed or direction (DVSA, 2014). For example, children running out from between parked cars or playing at the side of the road, blind and unmarked junctions or vehicles emerging from junctions would all constitute a hazard, as they would necessitate changing speed or direction. The hazards in the study were either static (e.g. junctions, roundabouts, parked vehicles, obstructions in the road, different types of crossings and traffic lights) or moving (e.g. pedestrians, cyclists, car drivers, motorcyclists, horse riders, large and hazardous vehicles, and disabled assistance vehicles). Participants were informed that a hazard was always present and that it did not occur in the first 10 seconds of each trial. Before the practice trials, participants were presented with information on what constitutes a driving hazard, including an example video. The two hazard responses were a keyboard press for the detection of the hazard and then a brief typed description of what the hazard was and its spatial location in the video. The videos were randomly allocated into any of the conditions including practice trials.

At the start of each trial a white fixation cross (0.6° x 0.6°) on a black background was presented for 1000 ms, followed by the video. Participants pressed the spacebar when they detected the hazard. This ended the trial and prompted presentation of the hazard response screen. If a participant did not make a keyboard response the trial ran for the duration of the video, followed by the hazard response screen.

In the no load condition participants only responded to the presentation of the hazard. In the low- and high load conditions, there was a concurrently-presented secondary task. In the low load condition, participants made a right foot pedal response to the presentation of the number 30 in the bottom right hand corner of the screen. In the high load condition participants made a 2AFC response according to whether the number was higher than 30 (31 through 40; left pedal) or lower (20 through 29; right pedal). The presentation of the number probe occurred on a random jitter (with a 1000, 2000 or 3000 ms inter-stimulus interval) from the start of the video until the participant responded to the hazard. Each number probe was presented for 3000 ms or until the participant made a response.

Load condition (no load, low load, high load) was blocked, with 20 trials in each block. The testing session also began with nine practice trials (with three practice trials of each condition), so participants completed a total of 69 trials. The presentation order of the three blocks was counterbalanced across participants. At the end of each block participants completed the Rating Scale Mental Effort (RSME; de Waard, 1996) to measure their subjective workload. As described in Chapter 1, the RSME is a unidimensional instrument to measure subjective mental workload, consisting of a 150 mm line marked with nine anchor points, each accompanied by a descriptive label and a corresponding number indicating a degree of effort (e.g. 0 = absolutely no effort, 120 = extreme effort). At the start of each block participants read a brief reminder of the task they were to complete. At the end of each block, as well as between trials, participants were given the opportunity to take a break. The entire experiment took 60 minutes to complete, including the practice trials.

Results

In all subsequent *F*-tests, the Greenhouse-Geisser correction was applied if Mauchly's test of Sphericity was significant, and I used the Benjamini-Hochberg false discovery rate procedure for all post-hoc tests. The false discovery rate was set at 0.1 for all multiple comparisons (Benjamini & Hochberg, 1995).

Data preparation

Workload measures.

Number probe responses were calculated from the first four responses per trial for each participant. The first four responses occurred on average before the mean hazard onset and were therefore unlikely to be influenced by responses to the hazard. This approach ensures that the mean responses are based on approximately the same number of trials for all participants across the low- and high load conditions. All responses with empty cells were removed (5.9%). Empty cells occurred when a trial was prematurely ended before the forth response (e.g. participants making a hazard detection response). Finally, all incorrect responses were removed (2.3%) for the RT analyses. The final dataset consisted of 4413 correct number probe responses (low load = 2258; high load = 2155), and overall 8.1% of the data were removed for the RT analyses.

Hazard detection measures.

In order to ensure that the hazards were clearly detectable across all the videos, data from videos with overall hazard response accuracies of less than 70% were removed from the analysis (17 videos, 21.72% of the data). Subsequently, response accuracy across all load conditions was high (M = 91%, SD =.29). All inaccurate responses were removed for the RT analyses (9.3%). In order to identify the earliest point at which a hazard could be detected, the fastest correct response was identified for each video and this time point was made the baseline against which all other hazard responses were measured (in 30 videos, the fastest response came from the participants; in the remaining videos, the fastest response came from an evaluator). The final dataset consisted of 1278 correct hazard responses, and overall 31.0% of the data were removed for the RT analysis.

Data Analysis

Workload Measures.

In order to examine the effectiveness of our load manipulation, a repeated measures ANOVA was conducted on the number probe RTs, using the within-subjects factor of load condition (low vs. high). There was a significant effect, $F(1, 29) = 348.26$, $p <.001$, $\eta_p^2 =.92$, indicating that RTs under low load (M = 570 ms, 95% CI = 530 – 608) were significantly faster than under high load (M = 978 ms, 95% CI = 922 – 1033), which confirms that the load manipulation was successful. This conclusion is supported by the accuracy analysis, which revealed significantly better performance under low load (M = 99.6%, 95% CI = 99.2 - 99.9) than under high load (M = 96.0%, 95% CI = 94.8 – 97.2), $F(1, 29) = 34.21$, $p <.001$, $\eta_p^2 =.54$.

The RSME subjective measure of workload was also tested using a repeated measures ANOVA with the within-subjects factor of load condition (no load, low load and high load). There was a significant effect of load condition, $F(2, 58) = 57.01$, $p <.001$, $\eta_p^2 =.66$ and post-hoc tests confirmed that each condition was different from the others (all p<.001; see Figure 2.1). These results indicate that the workload manipulation also worked as predicted from a subjective point of view.

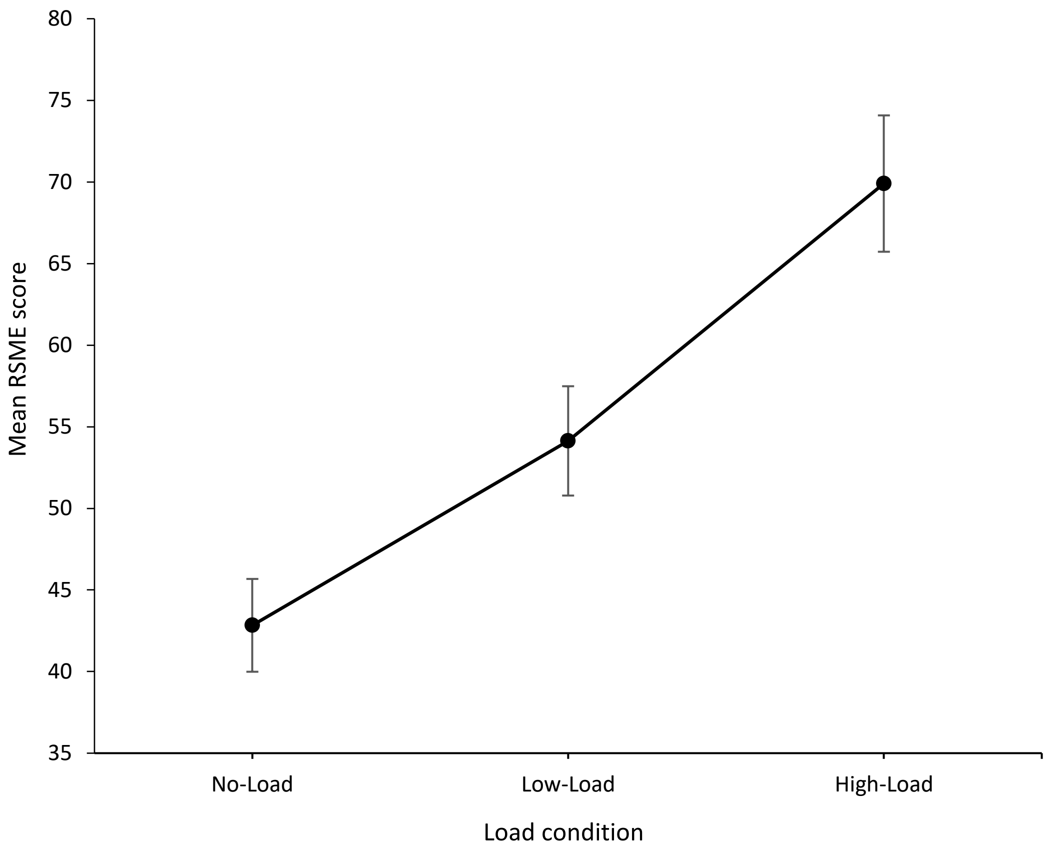

Figure 2.1. Mean RSME scores (± *95%CI*) as a function of load condition.

Hazard Detection Measures

Repeated measures ANOVAs were conducted on hazard detection RTs and accuracy using the within-subjects factor of load condition (no load, low load and high load). There was a significant effect of load condition in the RTs, $F(2, 58) = 3.45$, $p = .038$, $\eta_p^2 = .11$. Post-hoc tests identified a significant difference between no load and high load (p =.025), and a near-significant difference between no load and low load (p =.052), but not between low load and high load (p =.69; see Figure 2.2). These results indicate that the main effect was driven by the longer response times to the hazard in the no load condition than in the low- and high load conditions. Accuracy across the three load conditions was high (no load - M = 91%, 95% CI = 87.2 – 93.8; low load - M = 90%, 95% CI = 86.3 – 94.2; high load - M = 88%, 95% CI =

82.2 – 94.2), and did not vary significantly across load conditions, *F*(1.53, 44.49) = .426, *p* >.250, η_p^2 =.01.

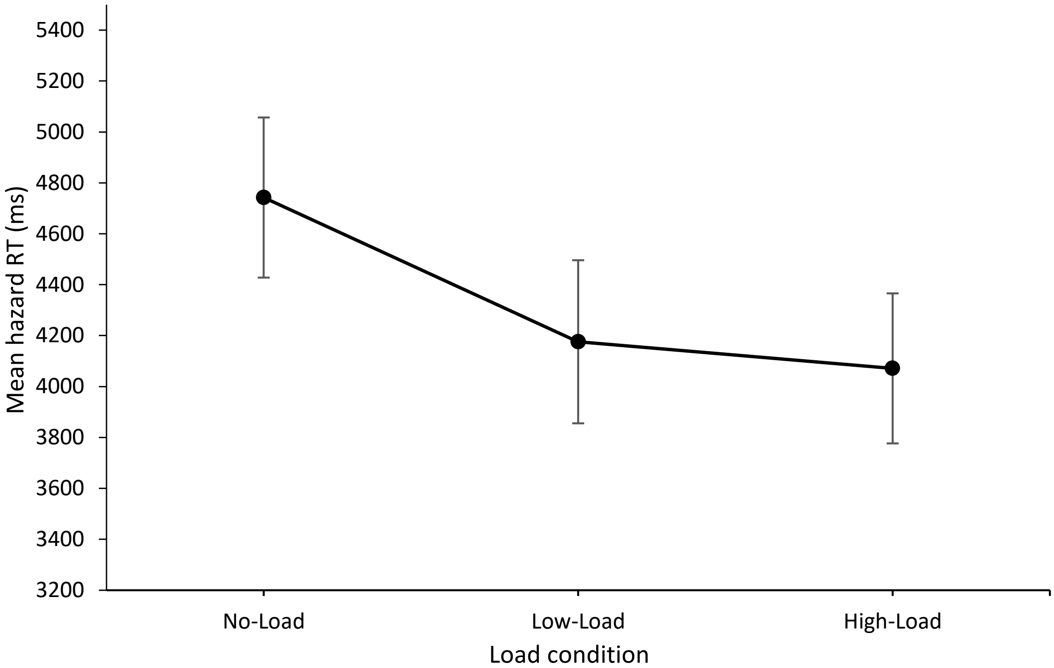

Figure 2.2. Mean correct response times (±95%CI) in milliseconds as a function of load condition.

Discussion

The present study demonstrated better hazard detection performance under dual (vs. single) task conditions. Response times for detecting driving hazards in realistic video stimuli were 560-680 ms faster when participants also responded to a numerical probe than when they completed the hazard detection task alone. This finding is extremely surprising given the wealth of studies demonstrating significant costs in dual (vs. single) task conditions.

However, the results may be argued to be in line with load theory, if one defines the hazards as ‘task-irrelevant’ stimuli. For example, according to the findings of Murphy and Greene (2017), increasing the executive control load (with the addition of the number task) should have increased attention to ‘task irrelevant’ items, perhaps explaining why the hazard was more effectively noticed in the presence (vs. absence) of the number task. However, given that the detection of the hazard was an explicit task requirement, this interpretation seems questionable. In addition, this explanation would predict better hazard detection under high load (vs. low load), but this pattern of results was not found.

The findings might instead relate to recent research examining the 'attentional boost' effect (Swallow & Jiang, 2010, 2012), which has demonstrated that detection of a target in one task can temporarily enhance processing of concurrently-presented (yet task-irrelevant) information. From this perspective, it is possible that superimposing a sequence of target stimuli onto a dynamic video stimulus could 'boost' processing of the dynamic stimulus, and this could explain the current findings. However, this possibility has never been tested. In fact, a key finding of the research has been that the attentional boost is temporally precise, with a duration of 100 ms before or after the target presentation (Swallow & Jiang, 2011, 2014). Thus, these findings may not relate directly to the attentional boost as it is currently defined but may instead raise the possibility of a similar but longer-lasting effect.

Interestingly, there was no difference in hazard perception RT between the low- and high load conditions. This non-difference could be because the high workload condition did not push the workload demands high enough. This could be an avenue for further research. However, recall that all three of the workload measures demonstrated significantly worse number task performance under high (vs. low) load, clearly indicating a reasonable level of difference in the demands imposed by the two different versions of the task. It is also possible that the presence of the number task simply increased participants' overall motivation, leading to improved performance on the hazard task whenever the number task was present (whether it had a high load or a low load). This seems unlikely, however, because hazard detection accuracy did not differ significantly across load conditions, as would be predicted by a motivation-based account.

A final plausible interpretation of the results of this study could relate to the Yerkes-Dodson law (Yerkes & Dodson, 1908) which describes a relation between arousal and performance. The key claim, which led to what later became known as the 'inverted-U hypothesis', is that performance is hindered at arousal levels that are either too low or too high (depending on the specifics of the task) with best performance seen at intermediate levels of arousal. In the context of task demand, this suggests that both underload and overload can negatively affect task performance. In the light of this research a more parsimonious explanation of the study results might suggest that the longer hazard response times in the no load condition are indicative of cognitive underload whereby the drivers' attentional resources were under engaged. By contrast, in the low and the high load conditions, attentional engagement may have approached an 'optimal' level, with no observed drop off as secondary task difficulty increased in primary task performance. This then raises the question as to whether the high load condition pushed the drivers to the edge of their attentional capacity limits, and therefore a future study with a very high load condition in the secondary task might show an increase in primary task RTs.

With respect to the driving literature, our findings are in line with the small number of previous studies showing decrements in hazard perception under automated (vs. manual) driving conditions (e.g. de Winter et al., 2014). Although our load manipulation was not designed in an attempt to mimic different levels of automation directly, the no load condition is somewhat analogous to a fully-automated vehicle, whereas the low- and high load conditions are approximately analogous to driving tasks that involve some degree of engagement. Thus, although fuller simulations are required before making firm conclusions, the results of this study suggest that keeping drivers engaged in some element of driving (e.g. lane control) is likely to improve hazard response performance by comparison with full automation.

The level of engagement that may be useful in this context is likely to vary from person to person according to their cognitive capacity. This raises the interesting question of whether a group of drivers with reduced cognitive capacity would show similar performance on this task as the young experienced drivers who took part in the current study. In Chapter 3, I examined this issue by applying the same paradigm used here to a driving population aged 65 years and over.

Chapter 3 – The Effects of Workload on the Perception of Driving Hazards in Older Drivers

Introduction

Older drivers, people typically aged 65+, are becoming an increasing discussion focus in the context of the effects of aging on driving safety. As the percentage of the general population aged over 65 increases, so do the number of older drivers. At present, there are approximately 4.1 million older drivers (65+ years) on UK roads (RAC, 2013), and by 2035 this number is predicted to increase to 21 million (AA, 2017). Furthermore, drivers aged 80+ involved in a traffic collision are six times more likely to be killed than drivers in their 40's (Clarke, Ward, & Truman, 2005; Clarke, Ward, Truman, & Bartle, 2009). The increase in older drivers, combined with an increasingly complex traffic and driving environment (e.g. high traffic volume, multiple lane junctions), could therefore increase the overall risk to the older driver, other road users and pedestrians. In fact, per mile driven, older drivers do have a higher crash risk but their overall crash risk per year is lower because they self-limit their driving exposure i.e. older drivers "compensate" by driving less (Department for Transport, 2016).

Numerous factors, such as age-related changes to visual attention, information processing, working memory (WM), eye movements, physical ability and eyesight can influence an older driver's competence to drive safely. However, the most frequent failures in older drivers involve visual search errors when turning right onto A-roads, and cognitive failures such as task fatigue and 'unintended acceleration' (Clarke et al., 2009). The caveat is that there is no specific age at which all drivers become unable to drive safely (RoSPA, 2010) because aging affects people differently.

With increasing levels of vehicle automation, the task of driving will become more focused on monitoring and responding to changes or failures in an autonomous system rather than the manual control of the car. What is currently unclear in the literature is how this change in driving will affect older drivers' behaviour, as they move from a highly practised skill (driving) to a relatively novel skill (system monitoring). Removing the driver partially or fully from being in the 'driving loop' (being involved in the complete task of driving) is likely to have different effects on cognitive function and information processing in older drivers than in younger experienced drivers (Endsley & Kaber, 1999; Gold, Dambock, Lorenz, & Bengler, 2013; Louw & Merat, 2017; Merat & Jamson, 2009). Drivers of autonomous vehicles spend significant amounts of time 'out of the loop', such that they are not involved in the task of driving except under certain conditions (e.g. inclement weather affecting systems or when there is a system failure). This makes it important to investigate their ability to make safe and appropriate responses when the autonomous system requires them to re-enter 'the loop' and regain control of the vehicle. The ability to re-enter 'the loop' and respond safely and appropriately could potentially be poorer in older drivers because of age related changes in cognitive and information processing. Therefore, it is crucial to quantify the effects of aging and reduction of task workload afforded by vehicular automation on safe driving performance. This will permit better assessment of whether the possible benefits of remaining engaged in some aspect of driving (staying in 'the loop') which were identified in younger experienced drivers in Chapter 2 might also apply to older drivers.

The key question this chapter will cover is: what are the specific effects of varied levels of task workload, used as a proxy for vehicle automation, on older drivers' situation awareness measured through the detection and perception of driving hazards? The chapter begins with a detailed literature review covering research from the cognitive and human factors domains, with particular focus on the way in which aging affects cognitive ability and information processing and how this in turn might affect older drivers' behaviour.

Visual Perception and Aging

For driving, two of the most basic and crucial aspects of visual perception are luminance (the amount of light cast on the retina) and motion perception. There is a wealth of research that shows that sensitivity to luminance, and to variations in luminance or contrast, declines with age, and that these changes can impact higher-level perceptual processing such as object detection (Seichepine et al., 2012). Crucially, older drivers' ability to perceive spatial changes in luminance or contrast seems likely to hinder their hazard perception ability, particularly under conditions of poor illuminance.

Motion perception also appears to decline with age. In a simple laboratory based study using moving dots, Trick and Silverman (1991) examined motion sensitivity in individuals between the ages of 25 to 80. They found a steady increase in motion thresholds with increasing age. In particular, the motion sensitivity thresholds for subjects aged 70+ were twice those of the 30-year-old subjects. Consequently, the ability of older drivers to detect motion may be compromised and, in a driving environment, this may affect their ability to accurately detect potential hazardous situations. Interestingly, it has been suggested by Bennett, Sekuler and Sekuler (2007), that age-related changes in motion perception may be due to reduced motion selectivity. They found that older participants (>70 yrs.) were significantly less sensitive to motion and were significantly less accurate at identifying the direction of movement compared to younger participants (23-50 yrs), and that the effect in older participants was not due to poorer WM or ability to use a computer. Their proposal was that increased internal visual noise in the visual systems of older participants leads the system to respond to a broad range of motion directions. These results suggest that changes in the tuning and firing rates of motion sensitive cells might partially account for the age-related declines in visual performance in older drivers. In sum, the decline in visual perception in older drivers is in part determined by changes to their motion perception ability. Given that the perception and response to motion is a significant part of the driving environment, the introduction of vehicle automation should reduce the impact of these age-related declines. However, automation is unlikely to negate the entire decline in motion perception in older drivers.

Another key aspect of changes in mid- to high-level visual processing in older drivers concerns optic flow – the perceived visual motion of objects as the observer moves relative to them. This is essential for detecting and avoiding collisions with surrounding objects. Studies have shown a small but significant decline in older participants' ability to determine an object's direction of locomotion (Warren, Blackwell, & Morris, 1989). This reduction in ability is suggestive of a decline in the use of optic flow information in high level visual processing of direction. The detection of an impending collision is an essential skill for any safe driver. If older participants are impaired at processing the direction of moving objects, then they are also likely to be impaired at detecting impending collisions. Andersen and Enriquez (2006) presented young and older participants with displays simulating approaching objects that would either collide or pass by the participant. The results showed that older participants had less sensitivity than younger observers in detecting collisions at increased speeds and shorter display durations. It is important to highlight that at low speeds young and old participants' ability to detect collisions on a linear path was not significantly different. It was only when the speed increased that the differences emerged. It has been found that at speeds of approximately 60 mph older drivers require on average an extra 2.5 secs of viewing the trajectory in order to have performance comparable to younger drivers (Andersen, Cisneros, Saidpour, & Atchley, 2000). This additional viewing time suggests that older drivers could have less time to initiate and complete a controlled response in order to avoid a potential collision. When they are moving, older drivers potentially find it more difficult to extract critical information from sources, such as the point of expansion and spatial location for approaching objects. It is highly probable that the impaired detection of an impending collision is a factor in the higher crash risk in older drivers.

In more recent research, Poulter and Wann (2013) examined why older drivers are more frequently involved in right of way collisions than younger drivers at road junctions, and their self-perceived competence of completing the junction manoeuvres. They used a task to measure drivers' ability to discriminate between different rates of looming presented by vehicles (car or motorbike) approaching at different speeds. Three age groups were used: 21–40 years; 50–70 years; 75+ years. The results showed a decrease in looming sensitivity to approach speed by between 2.8 mph for cars and 3.4 mph for motorbikes, for every decade that age increases. This result may reflect a decremental change in the central and peripheral visual fields as a result of age-related perceptual decline in older drivers. Additionally, the confidence ratings for all three age groups clearly showed evidence of an optimistic bias, with drivers in all age groups rating themselves as average or above average in regard to their ability to complete the task. Taken in total the results of this study show that older drivers make more right of way errors than younger drivers, were unaware they were making the errors and expressed confidence they were completing the manoeuvres safely.

In conclusion, it is clear from the perceptual evidence presented that there is no single underlying factor that influences the age-related changes in perceptual processing. Instead, the changes are driven by individual factors specific to the types of processing needed for a particular visual task. With all the studies presented the results should be interpreted with an element of caution. First, the studies are all laboratory-based and the validity of the results may not be as true if they were conducted in real life driving scenarios. Although laboratory-based studies can examine fundamental visual changes that will inevitably affect real-world vision, the data cannot always explain the changes seen in perception because of the complexity of the real-world visual environment compared with laboratory based studies. Second, the plasticity of the older participant's brain has not been examined in these studies. Therefore, if specific training was used it might improve the performance in these studies (see training section later for a more in-depth discussion).

Cognitive Changes and Aging

There are several cognitive changes that occur in older individuals, which affect their visual selective attention, information processing, working memory and eye movements. In this next section, I will briefly discuss these changes and their importance for safe driving.

Due to our limited processing capacity, we are unable to attend to all information available in the complex environments we live in (indeed, the driving environment could be argued to be one of the more complex of the environments that we encounter) (Broadbent, 1952, 1957; Deutsch & Deutsch, 1963). Using selective attention, our cognitive system can focus our limited cognitive resources on specific objects or locations of interest, and filter out any irrelevant information (Driver, 2001; Treisman, 1969; Treisman, 1964). This mechanism is essential in allowing us to respond effectively in the face of many competing stimuli, and any age-related impairments in selective attention are therefore likely to have significant impacts on the driving performance of older people.

Typically, older people are slower and less accurate than younger people when performing visual search tasks, suggesting an age-related decline in attentional functioning (Madden, 2007). However, this age-related decline in visual search performance appears to be influenced differently by top-down and bottom-up attentional factors. Specifically, top-down factors such as a person's expectations of a scene are relatively well-preserved as a function of aging (Madden & Langley, 2003; Wolfe, Butcher, Lee, & Hyle, 2003). Madden and Langley tested younger (19–27 yrs) and older (60–82 yrs) participants' performance during a letter search task in which a colour singleton was either noninformative regarding target location (baseline condition) or highly informative (guided condition). In the guided condition, both age groups exhibited a substantial decrease in RT to singleton targets, relative to the baseline condition, as well as an increase in RT to nonsingleton targets, indicating a similar pattern of top-down attentional effects in both age groups. However, in other studies aging has been found to impact top-down attention to some degree. For example, older people appear to be less successful than younger people at using a top-down attentional set (i.e., maintaining mental preparation) for avoiding attentional capture by a salient but task-irrelevant display item (Colcombe et al., 2003). Findings of this type indicate that some forms of top-down control exhibit an age-related decline, but older peoples' top-down attention exhibits some degree of preservation. In summary, research on age-related change in visual attention suggests that the decline seen in visual search performance with age is likely to be driven more by differences in bottom-up visual sensory processes than by differences in top-down processes, although aging effects are also seen in some aspects of executive processing related to attentional control.

In line with these claims, aging has also been argued to reduce perceptual capacity (which might be argued to reflect more of a bottom-up than a top-down influence). In a study that examined selective attention in older individuals, Maylor and Lavie (1998) asked twenty older participants (65+) to complete a two-choice visual search task in which perceptual load was manipulated by varying the number of letters in a circular display (set sizes of one, two, four, or six letters were used). Participants were instructed to ignore a distractor letter presented outside the circular display, which was either incompatible with the target or neutral. The results showed that the effect of perceptual load was greater in the older participants than in a younger (19-30 yrs.) comparison group. At low perceptual loads (smaller set size), the response- incongruent letters were a greater distraction than the congruent letters for older people, but these effects were reduced by comparison with the younger participants. This distractor effect also diminished quicker for older adults with increasing perceptual load. This means that lower levels of perceptual load were needed to reduce distractor interference in the older group compared with younger individuals. This is suggestive of lower perceptual load exhausting already reduced perceptual capacity for processing relevant targets in older individuals. This in turn leads to greater improvement in attentional selectivity, as a function of increasing perceptual load for older, not younger, individuals. In line with the claims of load theory, Madden and Langley (2003), suggest that when cognitive resources are limited the distraction from irrelevant display items generally decreases as perceptual load (display set size) increases. However, contrastingly, they observed that this perceptual load effect was not significantly different between young (18-24 yrs.) and older (60+) participants. They did find some evidence to suggest that there was an age-related decline in selective attention but that the decline was not due to reduced perceptual capacity in older participants but may be because of generalised cognitive slowing.

In sum, although there is not clear agreement from these studies on the question of whether aging reduces perceptual capacity. One interpretation of these results could be that the reduced workload associated with vehicle automation may negatively affect older drivers' selective attention because their distractibility has been shown to increase more than that of younger people in situations of reduced perceptual load. However, the outcomes of these studies need to be interpreted with care, as the results of laboratory-based studies with simple stimuli can be hard to generalise to real world driving environments.

Eye Movements During Driving and Aging

A highly important aspect of safe driving is the allocation of attention to those aspects of the driving environment that are appropriate for extracting information. This allocation of attention is observable through using eye movement data. The two basic eye movements are saccades (the movement of the eye) and fixations (the duration of time when the eyes remain fairly still and new information is acquired from the visual environment) (Rayner, 2009). The measurement of saccade amplitudes and velocities can be used to make inferences about cognitive processes. In laboratory based studies it has been consistently shown that older individuals exhibit longer saccade latencies (e.g. Klein, Fischer, Hartnegg, Heiss, & Roth, 2000) and decreased smooth-pursuit eye movements, older individuals make more saccades (rapid, ballistic movements of the eyes that abruptly change the point of fixation) rather than allowing the eyes to make slower tracking movements programmed to keep a moving stimulus on the fovea (e.g. Dowiasch, Marx, Einhäuser, & Bremmer, 2015), perhaps indicating less visual exploration of the environment than that which is shown by younger participants. For example, in a laboratory-based eye tracking study, Maltz and Shinar (1999), examined the eye movement performance of 10 younger (20-30 yrs.) and 10 older (62-80 yrs.) drivers whilst they viewed traffic scenes and responded to items in the scenes. Their results showed that older drivers spent significantly longer than the younger drivers searching the scenes, made more fixations on average (mean = 87 vs. 56 fixations) during the task and had shorter saccade amplitudes than the younger drivers. However, the average fixation durations were not significantly different between the groups. Fixation durations are argued to indicate the breadth and depth of cognitive processing, with longer fixation duration indicative of increased cognitive processing (but this can vary as a function of the task and the characteristics of the environment) (Torralba, Oliva, Castelhano, & Henderson, 2006). They also found that older drivers allocated a higher percentage of their time to searching smaller areas of the images, whereas younger participants searched the images more holistically. Taken as a whole, the results of this study suggest that older drivers need longer visual search times than younger drivers to extract the same information, even if they respond similarly on the task and have similar fixation durations.

However, not all research has identified significant differences between younger and older drivers' eye movement behaviour. Underwood, Phelps, Wright, van Loon and Galpin (2005), recorded the eye movements of older (60-75 years) and younger (30-45 years) drivers whilst they watched hazard perception driving videos. Participants were tasked with detecting other road users appearing on intersecting trajectories. Their results indicated that both groups were sensitive to attentional capture by the appearance of the hazard. Specifically, both driving groups detected the hazards presented in the study with a similar speed and accuracy, and also had similar scan paths towards the areas of the scene were the hazards occurred. However, the older drivers self-reported perceiving all the videos to be more hazardous in general than the younger drivers. In contrast to the studies already discussed, the results of this research suggest that there is no evidence of an age-related decline in hazard detection and visual search performance.

Another study that failed to find significant differences between younger and older drivers' hazard and scene perception was carried out by Borowsky, Shinar and Oron-Gilad (2010). 21 young, 19 experienced and 16 older drivers watched a number of hazard perception videos and were requested to identify hazardous situations. 75% of the videos contained a driving hazard, such as pedestrians crossing the road and approaching busy intersections. The remaining 25% of the videos contained no hazards and were used as control videos. There was no significant difference between experienced and older drivers' hazard perception detection rates, though older drivers had slightly slower RTs at busy junctions. The eye movement data showed that experienced and older drivers allocated attention to areas of the visual scene that contained information essential for the safe negotiation of busy junctions. In contrast, younger drivers typically fixated straight ahead; they neglected the valuable information that the experienced and older drivers fixated on in the scene. This study suggests that driving experience can improve a driver's situation awareness and guide attention towards potential hazards in the environment. Furthermore, age related factors under these circumstances have little effect on older drivers' ability to detect hazards.

However, research suggests that laboratory-based findings may differ from those elicited during real-world eye tracking. For example in a real world eye tracking study, 34 participants aged between 25 and 85 completed two real world tasks (walking down a hallway and visual tracking of an object) whilst having their eyes tracked. The older participants (60+) showed reduced saccade frequency and amplitude compared to younger participants (25-35 yrs.) similar to the findings of the laboratory-based studies. However, there was no difference in smooth pursuit eye movements between the older and younger individuals, which contrasts with the laboratory based results (Dowiasch et al., 2015). These results show that age-related eye movement changes as measured in the laboratory only partly resemble those that are seen in the real world.

There is evidence that training can improve the ability of older drivers to detect vital driving information. Rogé, Ndiaye, and Vienne (2014), used a simulated car driving task to examine the ability of 31 older drivers (65-75 years) to detect and respond to vulnerable road users, such as pedestrians and riders of two-wheeled vehicles. There were two groups: one group was given specific training to increase their useful field of vision, the other group were asked to just follow another vehicle. The useful field of view or perceptual span is the area around the point of fixation in which information is available for processing and interpretation by the visual system (Rayner, Castelhano, & Yang, 2009). It has been argued that the perceptual span deteriorates with age and may account for declining ability in older drivers to detect perceptual signals in a driving environment. This is particularly observed when conditions require the division of attention between central and peripheral tasks (Sekuler, Bennett, & Mamelak, 2000). The results showed that the trained group were able to detect pedestrians at a greater distance than the untrained drivers. The perceptual span training intervention used in this study was partially able to counter the reduction in elderly drivers' perceptual span, improving the visibility of pedestrians to older drivers. I would draw two particularly important points from this study: this form of training does have a lasting effect after 3 and 12 months, and secondly, regular training of this kind may delay or mitigate some of the cognitive decline in aging.

In conclusion, the results of the eye movement studies presented indicate that the process of aging can influence older drivers' allocation of visual attention to areas of interest in a scene. When specifically compared with younger drivers, older drivers make more saccades and have longer saccade latencies, and show decreased smooth-pursuit eye movements. In some studies, older drivers' attentional capture by hazards was no different to that of younger drivers, though their search of a visual scene may be less holistic. In addition, there is also evidence that the perceptual span of older drivers may decline, which could affect the processing of information in the periphery. Training interventions, however, may mitigate this decline. Overall, it seems that the effect of driving experience may mitigate some of the age related declines in cognition that were observed in hazard perception performance. However, it is important to bear in mind that the findings of real-world eye tracking studies have not always mirrored those of laboratory-based research (which makes up the bulk of the studies outlined here).

Dual Task Performance and Aging

The ability to process visual information concurrently from two or more competing sources of information is a highly important skill, especially in the context of driving, and changes due to aging are likely to have an impact on task performance (Ross, Dodson, Edwards, Ackerman, & Ball, 2012). For example, being able to process information about a vehicle's locomotion relative to yours, at the same time as monitoring the driving environment for other hazards involves the performance of two simultaneous tasks. Traditional dual task paradigms with young to middle-aged individuals typically show that reducing the number of concurrent tasks to be carried out tends to improve performance on the remaining tasks (Pashler, Johnston, & Ruthruff, 2001; Pashler, 1994; Watanabe & Funahashi, 2014). However, as highlighted in Chapter 2, I found the opposite pattern of results in a sample of the young experienced adult driving population. It is therefore difficult to predict how older drivers will perform on the same task. However, changes in cognition due to aging are likely to affect dual task performance more broadly (and therefore also related aspects of driving) as reviewed in the following paragraphs.

In a laboratory study by Glass et al. (2000), the effect of dual task performance in older individuals (60-70 yr.) was compared with younger individuals (18-26 yr.) using the psychological refractory period (PRP) paradigm. In a typical PRP study, two tasks are completed with longer or shorter stimulus onset asynchronies (SOAs) between the tasks. Typically, the shorter the SOA between tasks one and two, the longer the RT time in task two. The longer RT in the second task indicates that while the first stimulus is being processed, any other stimuli cannot be processed; the processing of the second stimulus is delayed, which slows down RT. Glass et al. found that participants' ability to coordinate the processing of two tasks did not decline with age. Longer RTs were found for all ages on the secondary task but this was not significantly different between the young and old individuals. Similar findings were presented by Hartley and Little (1999) who found no significant differences in dual task performance between young (M = 20 yrs.) and older (M = 73 yrs.) individuals. Specifically, they found that dual task inference increased with short SOAs, but did not vary between the age groups.

However, other research has observed larger dual task costs in older (vs. younger) participants. For example, in a study by Maquestiaux, Laguë-Beauvais, Ruthruff, Hartley, and Bherer (2010), 12 older individuals (65+) and 20 young individuals (18-25 yr.) were asked to learn to perform an auditory-vocal task, discriminate between the pitch of a tone presented for 150 ms (low vs. high pitch) and then at testing an additional concurrent unpractised visual task. The results showed that, compared to the younger individuals, older individuals were significantly poorer at performing two concurrent tasks rather than one single task. This was indicated by longer RTs and poorer accuracy to the low vs. high pitch task. In another study, Maquestiaux, Didierjean, Ruthruff, Chauvel and Hartley (2013) examined the effects of training older individuals to complete a simple novel dual task. Ten older (65+) individuals were asked to complete a dual task using the PRP paradigm over 12 separate sessions. The results showed that on the secondary task older individuals’ RTs were not significantly different to younger individuals (results from a separate study) (Maquestiaux, Hartley, & Bertsch, 2004). However, on the highly practiced primary task, RTs slowed on average on the secondary task compared to the first task by 485 ms for older individuals and only by 210 ms for younger individuals, indicating larger dual task interference in older (vs. younger) individuals. Potentially, the ability to automatise a task after practice is what could be driving changes in dual task performance in older individuals, rather than generalised cognitive slowing.

In conclusion, the results of dual task performance studies in older individuals are rather mixed despite well-controlled methods. It seems that under some conditions there is dual task interference to a greater degree for older drivers (vs. younger) and in other studies there are no age related differences. However, the effect of training on dual task performance is negligible, which would suggest that dual task deficits are likely to persist in older drivers, despite the additional experience that they are likely to have acquired.

Hazard Perception and Aging

Studies into hazard perception and aging have shown that age-related changes in perceptual and cognitive ability can increase RTs to detect a driving hazard and reduce hazard detection accuracy. There is evidence to suggest that drivers aged 65+ have poorer detection response rates and longer RTs than younger or middle aged drivers to hazardous situations on the roadway (Horswill et al., 2008; Horswill, Falconer, Pachana, Wetton, & Hill, 2015). This poor performance is often correlated with an increased risk of vehicular accidents (Wells, Tong, Sexton, Grayson, & Jones, 2008).

In general, drivers' abilities to detect driving hazards, and their RTs to these hazards, improve from the initial stages of driving up until 55 years of age. After 55, drivers' RTs to hazards begin to slow down (Quimby & Watts, 1981). The slowing of RTs after age 55 has been suggested to be driven by a number of different factors. Horswill et al. (2010) suggest that the increases in hazard perception RTs are driven by a combination of cognitive, sensory, and motor-response changes. These changes include a reduction in available cognitive resources, such as working memory deficits, which reduce the implementation of resource intensive hazard response strategies. In addition, impairments to the primary sense organs, such as vision and hearing, would affect the ability to rapidly identify hazards. A limitation to these studies is that the presence of neurodegenerative diseases in individuals was not controlled for at all. Rather than representing normal age-related changes in all older adults, the results could be influenced by the impact of these types of diseases.

In a video-based hazard perception study, Horswill et al. (2008), tested 118 drivers aged 65+ on their hazard perception abilities. A number of tests of cognitive ability, vision, and simple RTs were carried out as well. Their hypothesis was that factors other than cognitive ability could account for the increase in RT to detect driving hazards. The results suggest that hazard perception RTs do significantly increase with age. Specifically, it can be surmised that a decrease in contrast sensitivity, a narrower perceptual span and slowing in simple RTs can account for the variance in hazard perception, independent of individual differences in age. However, it is unclear from this conclusion which of the three changes accounts for the largest proportion of the variance in RT in this study. A clearer way to address the question of older driver's hazard perception ability would be to in a series of studies control contrast sensitivity using appropriate visual filters, to have a control group of regular adult drivers to compare the results to, and to use a variable size visual mask to understand the changes to their perceptual span.

Change detection is a central part of hazard detection. Change blindness is a cognitive phenomenon where people are surprisingly poor at spotting major changes between two images, when the original and altered image are quickly alternated (Simons & Rensink, 2005). Although hazard detection and change blindness are similar in that they are both perceptual errors by the visual system e.g. failing to look and looking but failing to see, the former can be explained in regard to poor visual scene scanning, for example failing to check the blind spot. The latter is harder to understand – how can a driver actively search a visual scene, yet not process visible and vital changes in the visual information? The use of Simons-type methodologies may help elucidate this latter question. In a simple laboratory based study using a change blindness paradigm, Caird, Edwards, Creaser and Horrey (2005), presented young (18-25 years), middle-aged (26-64 years), old (65-73 years), and older (74+ years) drivers with 36 pictures of junctions with one object manipulated in the scene. The object in the picture could be a pedestrian, vehicle, sign, or traffic control system and, using a flicker paradigm, the scene would change every 5 or 8 seconds. Participants were asked to make a decision about whether or not it would be safe to proceed through a junction, turn left, or turn right. The results showed that young and middle-aged drivers made significantly more correct decisions than did old and older drivers. In particular, both groups of older drivers were less likely to detect a hazard, and were particularly less likely to detect the presence of pedestrians. Overall, it seems that older drivers are more susceptible to missing changes in a busy driving environment and, in this example, missing at risk potential targets such as pedestrians. It could be argued that partial vehicle automation (e.g. level 3) could improve older drivers' hazard perception by allowing them to attend more to the driving environment and requiring less switching of focus between the roadway ahead and behind, as well as the internal monitoring of systems. Reducing the number of tasks older drivers complete during autonomous driving (vs. non-autonomous driving) should free up attentional resources, which are already reduced in old age, thus allowing them to use this resource capacity to better carry out hazard monitoring and readying for resumption of control.

Road signs and signals are important tools to convey critical information to drivers for safe driving and efficient navigation. Both the physical and cognitive changes associated with aging will affect the acquisition and processing of the information present on these road signs. For 85 American road signs, Dewar, Kline, Scheiber and Swanson (1997) measured the legibility distance under night and night-with-glare conditions, as well as RT to correctly identify the sign and conspicuity (how easy it is to see the sign). They observed that increasing driver age was associated with lower comprehension levels and legibility distances for the road signs, under all test conditions, compared to younger drivers. In addition, compared to younger drivers, older drivers had longer RTs in order to make the appropriate response to a sign. Their search times were also greater when the conspicuity of the sign was reduced, compared to the younger drivers.

In conclusion, ageing affects one of the key aspects of situation awareness, hazard perception, but it is not clear-cut as to which factors (sensory, cognitive, perceptual or motor) have the biggest effect. In general, older drivers have fast RTs to hazards up until around the age of 55. After this, point RTs and accuracy in hazard detection drop off. Whilst it is not entirely clear, these longer RTs and poor accuracy may in part be due changes in cognitive process and changes in vision. Compared to younger drivers, older drivers seem to miss certain types of hazards more than others (e.g. pedestrians). However, the effect of experience in responding to and perceiving hazards may partially mitigate some of the negative effects of aging.

Perceptual and Visual Training

The most effective way to reduce the crash risk among older drivers is to encourage individuals to limit their driving exposure or even cease driving altogether. However, this should be seen as a strategy of last resort, as the negative impact of reducing or stopping driving has been demonstrated to increase isolation, loneliness and increase the risk of depression in older drivers (Marottoli et al., 1997). A more nuanced approach is to encourage older drivers to find alternative strategies or partake in additional training to counteract their driving deficiencies. However, it is a point of contention as to whether a training intervention for older drivers is likely to be effective. Older drivers are likely to have many years of driving experience across a large variety of roadways and traffic scenarios. Consequently, it could be argued that any specific hazard perception intervention is unlikely to alter significantly their driving behaviour and driving habits built up over time. A combination of hazard perception, visual search and driving behaviour interventions may have more of an effect on older drivers' behaviour than hazard perception training on its own.

There is some evidence to suggest that driving training interventions can improve hazard perception and reduce RT, even in highly experienced drivers.

McKenna and Crick (1994) found that experienced police drivers who had received hazard perception training (e.g. in understanding how hazards develop, where hazards commonly appear and how to respond to hazards) were significantly faster at responding to hazards than drivers with equivalently high levels of experience, but no specific training. Importantly, they found that when experienced drivers were given hazard training they could improve their RTs in video simulations. This means that, even after years of driving experience and hundreds of thousands of miles driven, drivers can benefit from training interventions.

There are a number of studies on hazard perception training interventions in older drivers showing that the training can improve hazard perception ability. These findings highlight the importance of careful counterbalancing of video and trial order in research in this area (including the present study), in order to ensure that these types of training-related improvements do not confound the central experimental manipulations. Horswill et al. (2010) used a video-based hazard perception training intervention on 28 older drivers (65-94 yrs.). These 28 drivers were split into two groups: one group received hazard perception training; the other group did not and acted as the control group. Hazard perception RTs were obtained for both groups before and after the training intervention was implemented. The results show that the group that received training had significantly shorter RTs than the untrained group, with RTs in the order of 500-1000 ms quicker from both groups' pre-training baseline RT. There was no significant interaction between the group and their pre-training baseline RTs. This 500-1000 ms speeding up of RTs due to the intervention could make a significant difference in the likelihood of experiencing a crash, by providing 8-9 metres more road to make a response if travelling at 35-40 mph. However, it should be noted that these results may not translate to real world driving scenarios and that the effects of the training may not have any long lasting effects on hazard perception ability.

Nevertheless, the question of whether a lasting effect can be achieved was addressed in a recent study by Horswill et al. (2015). 75 drivers aged 65 and over received either a 35 minute video-based hazard perception training intervention or no intervention at all. The results showed that this brief training period led to improvements in hazard detection accuracy and RT, and that these benefits persisted for approximately 1 month and 3 months after the intervention, without a significant decay in the training effect over this period.

In recent years, heads-up displays (HUDs) are increasingly used to display useful driving data within the car. With this in mind, studies have been conducted to assess the effectiveness of augmented reality (AR) technology to mitigate accident risk by directing older drivers' attention to potential hazards in the driving environment. One such study by Schall et al. (2012), investigated whether AR cueing improved or interfered with hazard perception in older drivers. Using a sample of 20 older drivers (65+) they used a number of AR cues to direct attention to potential roadside hazards (e.g. pedestrians) in a driving simulator. The results showed that older drivers' detection of low visibility driving hazards was improved compared to a baseline measure. AR cueing did not interfere with the detection of non-hazardous objects in the driving environment and, importantly, did not impair the ability to maintain a safe headway to the vehicles in front in older drivers. However, a younger experienced driving comparison group was not utilised in this study, therefore it is hard to determine whether the pattern of results observed relates specifically to older drivers or could in fact apply to drivers of all ages.

Eye scanning training is designed to teach drivers to scan the environment more effectively and efficiently, in part by observing the eye movements of very experienced people, and is similar to perceptual span training which aims to widen the focal point of visual attention. Pollatsek, Romoser and Fisher (2012) observed that for crashes at road junctions older drivers are at greater risk of accident and this may, in part, be due to insufficient scanning of the driving environment, rather than deteriorating physical or cognitive capabilities. They observed that older drivers seem to develop a search strategy that focuses attention towards regions immediately ahead of their vehicle at the cost of scanning peripheral junction locations likely to contain potential hazards. Two groups of older drivers (65+) were allocated to either an active eye scanning training intervention or a passive intervention. The eye scanning training group saw a video replay of their behaviour at junctions and most of the older drivers recognised their failures without prompting. They were also evaluated in a driving simulator, shown feedback of their behaviour at the virtual junctions, and then allowed to practice appropriate scanning behaviours on the simulator. In the passive group the participants were given 30 to 40 minutes of instruction, which included coaching about where to look at junctions and why less careful scanning of areas from which potential hazards might emerge was an important cause of many crashes for older drivers. The results showed that both groups' eye scanning improved but importantly the active intervention group's performance increased dramatically, to a level that was indistinguishable from the performance of younger experienced drivers (from a previous study). The effectiveness of the eye scanning training was shown to have endured at both 3 and 12 months. However, what is not known is if these training interventions actually reduced the older driver's crash risk in practice.

Further research has also bolstered the claims of training interventions to improve visual search performance. Lavallière, Simoneau, Tremblay, Laurendeau and Teasdale (2012), evaluated the effectiveness of video-based feedback training in 22 older drivers. Ten older drivers (65+) received training on a number of driving performance measures (e.g. hazard perception, headway distance) and driver-specific feedback on their search behaviour (e.g. how they were performing). The control group of 12 older drivers (65+) received similar training but received no feedback. The results showed that after training the control group showed no increase in the frequency of the visual inspection of three regions of interests (rear view and left side mirrors, and blind spot). However, importantly, the feedback group significantly increased the frequency of visual inspection of the three interest regions compared to the control group. In sum, the results of these studies suggest that eye scanning training with specific driver feedback helped older drivers to improve their visual inspection strategies. Furthermore, the beneficial effects of this training can endure for up to a year.

In conclusion, the effectiveness of training programmes to improve older drivers' performance on important driving tasks seems to be generally positive. Numerous types of training with elderly drivers have been shown to reduce their

RTs to hazards and improve their hazard detection, and crucially, the effects have been shown to endure. The results of studies like these show that driving training can be effective for older drivers but it remains to be seen if training interventions can actually reduce the risk of road collisions.

Overall Conclusions

The key point to take from the research into older individuals and their driving behaviour is that there are numerous factors that can influence an older driver's competence to drive safely but there is no one predominant single factor. The perceptual processing changes due to aging are varied and well-established in the literature. Cognitive accounts of changes in older individuals would suggest that at very low loads older individuals are more distractible, but with small increases in load their capacity is exhausted, leaving them less susceptible than younger individuals to distraction at intermediate and higher levels of load. In addition, older people are more likely to miss changes in the environment from moment to moment. From an eye movement perspective, the findings are not entirely clear-cut, with some weak evidence suggesting that age can influence the patterns and types of eye movements made. Some evidence does suggest that older individuals do have changes to their spatial attention as the product of age, but this is not always the case. There is also a mixed pattern of evidence from dual task paradigms investigating information processing in the elderly. However, there does seem to be some effect of dual task workload on the ability of older individuals to complete laboratory-based studies with similar RTs in the primary task, as younger individuals. Older individuals' hazard perception abilities (longer RTs, poorer accuracy) seem to be compromised compared to experienced younger drivers, but this drop off may only start to occur after the age 55. Finally, training interventions to improve older drivers' hazard perception, perceptual span and visual search performance do on the whole seem to improve their performance on key measures (e.g. visual search strategies, hazard perception RTs).

In the present study, a group of older drivers completed the same task as had been used with younger experienced drivers in Chapter 2. Given the mixed pattern of findings concerning whether older drivers are more susceptible to dual task effects, along with the fact that the younger experienced drivers in Chapter 2 exhibited an unexpected improvement in performance under dual (vs. single) task conditions, it is hard make clear predictions. Nevertheless, based on the assumption that cognitive capacity does on the whole decline with age, I predicted that the older driving cohort would show an increase in hazard detection RTs as the dual task workload increased.

Method

Participants

Thirty older drivers were recruited for the study through TRL (mean age = 75.2, SD = 4.3, 20 males). The sample size was derived from prior research, constituting a typical size for studies in this area (e.g. Galpin, Underwood, & Crundall, 2009). 26 participants were right-handed. All had a full UK driving licence and a mean driving experience of 53 years (SD = 6.5). The average miles driven per year was 6542 (SD = 3150). Participants were naïve to the purpose of the experiment, and all had previously taken part in driving experiments at TRL. None of the older drivers had taken the hazard perception test as part of their driving theory test. All participants had normal or corrected-to-normal visual acuity, gave informed consent before participating, and were compensated with £20 for participation in the study. All procedures were reviewed and approved by the Departmental Ethics Committee.

Stimuli and Design

The apparatus, stimuli, design and procedure were all exactly as described in Chapter 2.

Results

As in Chapter 2, in all subsequent *F*-tests, the Greenhouse-Geisser correction was applied if Mauchly's test of Sphericity was significant, and the Benjamini-Hochberg false discovery rate procedure was used for all post-hoc tests, with the false discovery rate set at 0.1 for all multiple comparisons (Benjamini & Hochberg, 1995).

Data Preparation

Workload measures.

As in Chapter 2, number probe responses were calculated from the first four responses per trial for each participant. All responses with empty cells were removed (4.5%). Empty cells occurred when a trial was prematurely ended before the forth response. Finally, all incorrect responses were removed (8.2%) for the RT analyses. The final dataset consisted of 4188 correct number probe responses (low load = 2077; high load = 2111), and overall 12.7% of the data were removed for the RT analyses.

Hazard detection measures.

As in Chapter 2, in order to ensure that the hazards were clearly detectable across all the videos, data from videos with overall hazard response accuracies of less than 70% were removed from the analysis (23 videos, 33.1% of the data). Note that this procedure was carried out separately for this data set (rather than basing these video eliminations on the data from Chapter 2) in order to ensure that all hazard videos used contained hazards that were clearly detectable for all participants tested. Interestingly, slightly more videos needed removal for this older population than had been the case for the younger experienced drivers in Chapter 2 (where only 17 videos were removed). Following the removal of the low accuracy videos, response accuracy across all load conditions was high (M = 81%, SD =.40). All inaccurate responses were removed for the RT analyses (18.6%). Once again, as in Chapter 2, in order to identify the earliest point at which a hazard could be detected, the fastest correct response was identified for each video and this time point was made the baseline against which all other hazard responses were measured (in seven videos the fastest response came from the participants; in the remaining videos, the fastest response came from an evaluator or previous experiment benchmark). The final dataset consisted of 980 correct hazard responses, and overall 45.6% of the data were removed for the RT analysis.

Data Analysis

Workload Measures.

In order to examine the effectiveness of the load manipulation, a repeated measures ANOVA was conducted on the number probe RTs, using the within-subjects factor of load condition (low vs. high). There was a significant effect, $F(1, 29) = 387.23$, $p < .001$, $\eta_p^2 = .93$, indicating that RTs under low load (M 700 ms, 95% CI = 648 – 751) were significantly faster than under high load (M = 1076 ms, 95% CI = 1034 – 1118), which confirms that the load manipulation was successful. This conclusion is supported by the accuracy analysis, which revealed significantly better performance under low load (M = 86.7%, 95% CI = 82.5 – 90.9) than under high load (M = 82.1%, 95% CI = 76.4 – 87.8), $F(1, 29) = 5.72$, $p = .023$, $\eta_p^2 = .17$.

The RSME measure of workload was also tested using a repeated measures ANOVA with the within-subjects factor of load condition (no load, low load and high load). There was a significant effect of load condition, $F(1.2, 36.1) = 15.22$, $p < .001$, $\eta_p^2 = .34$ and post-hoc tests confirmed that each condition was different from the others (all $p < .001$; see Figure 3.1). These results indicate that the workload manipulation also worked as predicted from a subjective point of view.

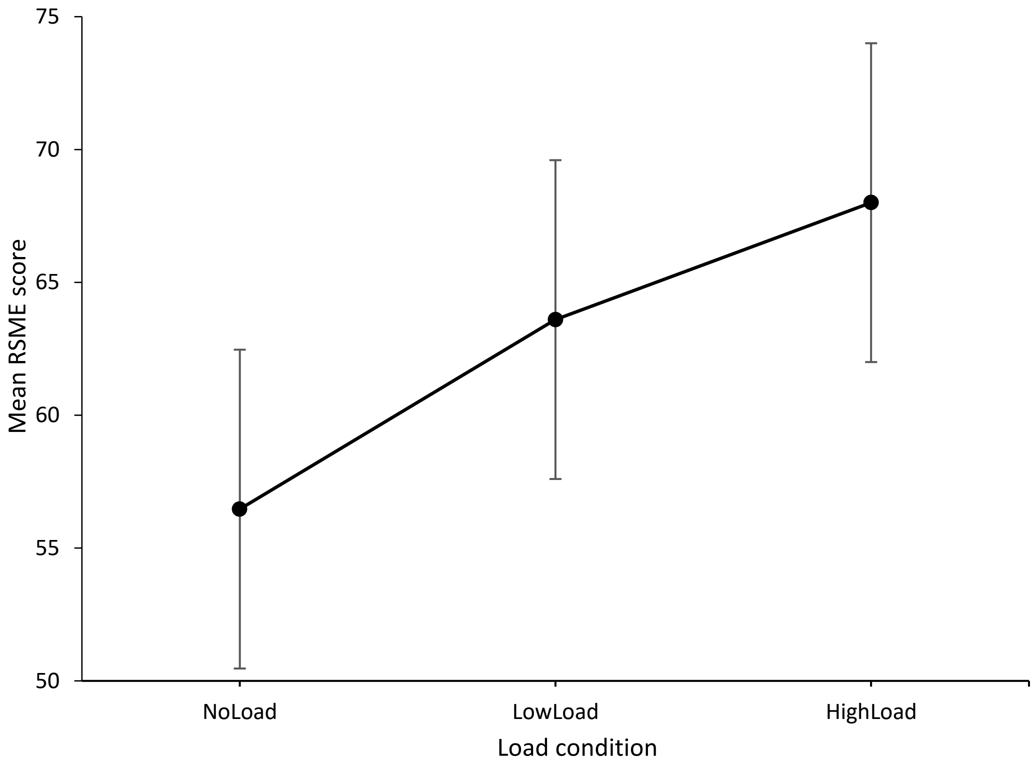

Figure 3.1. Mean RSME scores (± 95% CI) as a function of load condition.

Hazard Detection Measures.

Two repeated measures ANOVAs were conducted on hazard detection RTs and accuracy using the within-subjects factor of load condition (no load, low load and high load). There was a significant effect of load condition on the RTs, *F*(1.62, 46.56) = 3.48, *p* =.049, η_p^2 =.11. Post-hoc tests identified a significant difference between no load and high load (*p* =.001), but no significant difference between no load and low load (*p* =.25) or between low load and high load (*p* =.23; see Figure 3.2). There was also a significant effect of load condition in the accuracy data, *F*(2, 58) = 3.93, *p* =.025, η_p^2 =.12. Post-hoc tests identified a significant difference between low load (M = 77.6%, 95% CI = 69.7 – 85.5) and high load (M = 85.7%, 95% CI = 79.9 – 91.4, *p* =.017). Note that this effect is in the opposite direction to that seen in the RTs, with worse performance under low (vs. high) load. There was no significant difference between the no load (M = 82.2%, 95% CI = 76.2 – 88.1) and low load conditions (*p* =.11), or between no load and high load (*p* =.19).

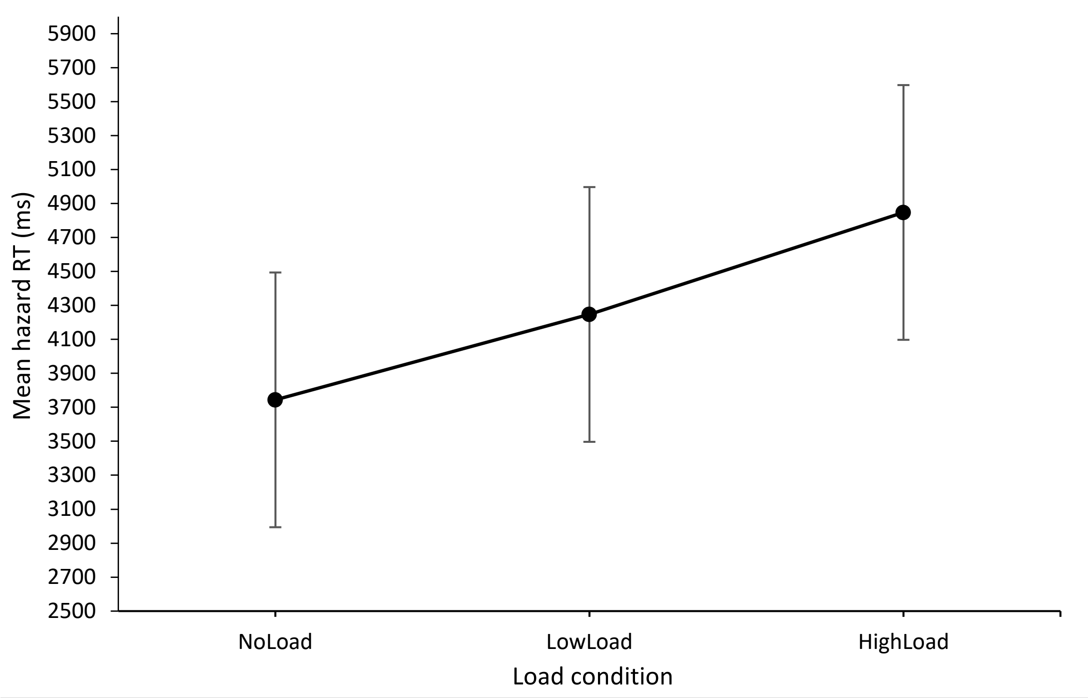

Figure 3.2. Mean correct response times (±95% CI) in milliseconds as a function of load condition.

Because the accuracy data revealed a pattern that was opposite to that of the RT data, participants' hazard detection RTs and accuracy data were converted to inverse efficiency (IE) scores in order to account for any potential speed–accuracy trade-offs that might have been present. An IE score is calculated by dividing each participant's mean RT for each condition by the proportion of correct responses for that condition, so lower IE scores indicate better task performance. A repeated measures ANOVA was conducted on IE scores using the within-subjects factor of load condition (no load, low load and high load). The analysis revealed no significant main effect of load condition. Participants' hazard detection performance was not significantly different between the no load (M = 4842, 95% CI = 3557 - 6127), low load (M = 8557, 95% CI = 1640 - 15474) and high load conditions (M = 6347, 95% CI = 4200 - 8495), $F(1.1, 30.9) = .96$, $p = .34$, $\eta_p^2 = .03$.

Discussion

The results of the present study demonstrate that older drivers' RTs for detecting driving hazards in realistic video stimuli were on average 1100ms longer when they also completed a high load number probe task than when they completed the hazard detection task alone. This finding is in line with previous research on dual task interference, which typically finds significant costs in dual (vs. single) task conditions (Ettwig & Bronkhorst, 2015; Watanabe & Funahashi, 2014). Interestingly, however, hazard response accuracy was significantly higher, by 8%, in the high load condition (vs. low load), and not significantly different between the no- and high load or no- and low load conditions, as might be predicted from a dual task interference account. This raises the potential of a possible trade-off between the accuracy and RTs in the hazard detection task for older drivers. The results of the IE analysis (in which no differences were found between conditions) indicate that there might indeed have been a speed accuracy trade-off. The longer overall RTs as task difficulty increased were accompanied by an increase in accuracy. This could be suggestive of a strategy change by older drivers, such that they sacrifice speed as task load increases, in order to achieve good accuracy. Chapter 5 will return to this question, through the use of direct statistical comparisons between the results of Chapters 2 and 3, as well as a more detailed analysis of the accuracy data.

In the meantime, what can be drawn from the results of this study is that older drivers' hazard perception follows a pattern that shows that as task workload increases their hazard detection RT performance slows but their hazard detection accuracy increases. This could have implications for older drivers when operating an autonomous car. Specifically, the reduction in task demand as a result of increasing levels of automation could benefit older people's driving performance by allowing the driver to detect and respond to driving hazards quicker, however these improvements in response speed may be accompanied by reductions in the accuracy of hazard identification.

Chapter 4 – The Effects of Workload on Hazard Perception in Novice Drivers

Introduction

Having examined the impact of a dual task load on hazard perception in younger experienced drivers (Chapter 2) as well as older drivers (Chapter 3), the present chapter investigates performance on the same task by young novice drivers. It is particularly important to consider performance in this group because in the UK young novice drivers (17-21 years old) make up only 1.5% of all licence holders (DVLA, 2015) yet they are involved in 9% of all fatal or serious road accidents (Department for Transport, 2015). On average a driver aged 16-19 years old is three times more likely to crash than a young experienced driver (40-49 yrs. old) (Department for Transport, 2015), with almost a quarter (23%) of 18-24 year olds involved in a crash within the two year probationary period after passing their driving test (Butcher, 2016). A key contributing factor in 55 fatal accidents in 2015 and 4,483 personal injury accidents was cited as learner or inexperienced driver/rider (Butcher, 2016; Department for Transport, 2015). In the US the leading cause of death in teenagers is vehicular accidents (Centers for Disease Control and Prevention, 2017). Specifically, figures from 2014 in the US show that 2,333 drivers aged 16-19 years old were killed and more than 221,313 were treated in hospital for motor vehicles crashes (Centers for Disease Control and Prevention, 2015).

In this section, I will discuss evidence to suggest that young novice drivers differ from younger experienced drivers in patterns of visual attention, ability to manage task workload, executive function, and hazard perception. These differences may explain the young novice drivers' higher crash rates and higher driving risk. They also inform the predictions concerning the outcome of the present study, as I will discuss at the end of the introduction section.

Young Novice Drivers' Visual Attention

As explained in previous chapters, attending to task relevant information in a timely manner is an important component of safe driving (Sivak, 1996). However, it seems that for young novice drivers there can often be deficiencies in the patterns of visual attention and information processing that increases their driving risk. In a literature review, Underwood (2007) found that novice drivers show different visual search behaviour to experienced drivers. Underwood found that when drivers scan the road around them, differences are observed as function of driving experience and training, with experienced drivers increasing their visual scanning on roadways of increasing complexity. However, young novice drivers do not show this sensitivity to road complexity, suggesting that they fail to attend to potential dangers involving the behaviour of other road users. Underwood suggests that that the driver's understanding of the task develops with experience, such that roadways that demand increased monitoring (e.g. interweaving traffic on a multi- lane roadway) receive more extensive scanning than simpler roads (e.g. light traffic on a straight rural road).

The most predominant visual strategy used by both novice and experienced drivers is to fixate straight ahead at the position in the roadway where their vehicles will be in the next few seconds. This focus of expansion (FOE) is generally equal to the direction of the vehicle, particularly on straight roads, and this is the fixation location preferred by experienced and novice drivers (Mourant & Rockwell, 1972; Underwood, 2007; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003). Outside of this central fixation point, most other fixations are observed to fall on the horizontal axis to the left and the right of the FOE, as this is the area where task relevant information will appear (e.g. road signs, pedestrians, parked cars and other potential hazards). Thus during driving almost all fixations fall in an observed elliptical window with very few fixations falling in the vertical axis in an experienced driver (Chapman & Underwood, 1998; Crundall, Underwood, & Chapman, 1999). However, young novice drivers show a more varied distribution of fixations particularly in the vertical axis (Crundall & Underwood, 1998; Mourant & Rockwell, 1972), which indicates that scanning of the horizontal visual axis is a skill developed through experience, particularly of where to expect hazards or task relevant information.

Using a driving simulator and eye tracking methodology to examine the navigation techniques and eye movement behaviour changes due to age, Scott, Hall, Litchfield, and Westwood (2013) compared the fixation patterns, including the number of fixations, gaze frequency (the frequency of making return fixations to areas of interest) and gaze duration, of three driver groups (young novice, young experienced and older experienced) during simulated driving through a series of complex junction manoeuvres. The study was motivated by the finding from accident statistics that older drivers and young novice drivers have problems negotiating certain road junctions. The fixation patterns of the three driver groups were compared during gap selection at a right turn junction (left turn in countries driving on the right), which is where higher levels of crashes are typically seen, particularly in older and novice drivers (McGwin, Jr & Brown, 1999). 14 novice drivers (mean age 20), 14 younger experienced drivers (mean age 23) and 14 older experienced drivers (mean age 66) took part in the study. Drivers were instructed to make a right turn manoeuvre in their own time, then stop at a predefined point on a straight section of road. The results showed that when scanning the junction, younger experienced drivers distributed their gaze more evenly across all areas, whereas older and novice drivers made more sweeping transitions (fixating in a less uniform and more chaotic manner). Additionally, during the scanning or preparation to move phase all drivers adopted a preview strategy in which they predominately searched between the middle and far areas to the left and the right of the junction. However, both younger driver groups (young novice, young experienced) also showed gaze transitions from middle right to near right and near left to middle left, whereas older drivers did not adopt this strategy. It is possible that this strategy difference between young and old drivers is driven by a decrease in cognitive capacity due to aging. In the decision phase (i.e. deciding to make the manoeuvre) the use of a preview strategy was less evident in the older experienced group compared to the younger groups. It is possible that response preparation requirements of the decision phase impact on older drivers' ability to maintain a preview strategy to complete the manoeuvre.

The results of this study indicate that young novice drivers share some commonalities in visual sampling behaviour with younger experienced drivers. However, they are still over represented in the crash and driver fatality statistics. This in part is likely to be due to factors beyond visual scanning behaviour such as poor hazard perception, low executive function or cognitive underload (i.e.

situations that involve very low task demands, which can negatively impact the performance of the person carrying out the task by increasing RTs and decreasing accuracy to the task). A note of caution should be taken with these results as they indicate visual scanning behaviour during a very specific scenario and the pattern of results may not extrapolate out into more general driving behaviour.

In a seminal study on the effect of driving experience on visual attention and search, Mourant and Rockwell (1972) showed some key differences in the visual search and information acquisition strategies of novice and experienced drivers. They tested six novice drivers (all male, aged 16-17) and four experienced drivers (all male aged 21-43, with an average mileage of 8000/year). Participants drove a suburban and motorway route while their driving behaviour (vehicle velocity, acceleration, brake and accelerator usage) and eye movements (fixation points, pursuit eye movements and blinks) were recorded. Novice drivers were found to fixate in a smaller central area in the visual scene than experienced drivers, looking close to the front of the vehicle and sampling their mirrors less frequently. These patterns are suggestive of a less well-developed process of visual acquisition in the novice drivers. They also showed a right side horizontal axis bias compared with the experienced drivers. This is indicative of searching for the edge of the road, which is then used to guide lane positioning. This suggests that the novice drivers were not automatized in their driving behaviour and were potentially using more cognitive resources to maintain a safe lane position than the experienced drivers (for whom the lane alignment is likely to have been a more automatic cognitive process). Additionally, the novice drivers made more pursuit eye movements (i.e. fixated on tracking elements of the visual scene) than the experienced drivers. This appears to be indicative of a strategy that fixates the eyes closer to the front of the vehicle and its relative lane position, with the eyes following the road markings as the vehicle moves. By contrast, the experienced drivers exhibited fewer pursuit eye movements as their driving style is likely better developed, such that they can sample areas of the scene on the horizontal axis that contain task relevant information, allowing them safely to navigate different roadway environments.

In sum, the results of Mourant and Rockwell's (1972) study show that the search and fixation strategies of novice drivers might impair their ability to detect hazardous road situations. However, the novice drivers whilst taking part in this study were all entered in to a driver training program between each route they drove. It is possible therefore that the driver training element was influencing their driving behaviour because they were driving cognitively demanding routes whilst also learning the demanding task of driving a vehicle. Therefore, it is difficult to untangle the effects of task workload in this study, because the driving and learning to drive elements are conflated together.

However, some research has examined the effect of task load more directly and using larger participant numbers. In a key paper in this field, Crundall and Underwood (1998) examined the visual search and information acquisition strategies of novice and experienced drivers under different levels of cognitive load imposed by different types of road. Sixteen experienced drivers (with mean experience of 9 years) and 16 novice drivers (with mean experience of 2 months), all of whom held full licences, were asked to drive a pre-specified route for 20 minutes. During this route they drove through three distinct areas: first a rural, single-lane carriageway; second, a suburban road through a small village which contained some shops, parked cars and zebra crossings; and third, a dual carriageway with two lanes of forward-moving traffic and traffic merging from the left. The latter two were selected for inclusion in the test route because they placed the driver under a higher level of visual cognitive load than the rural road. During the 20 minutes’ drive the participants had their eyes tracked and were instructed to drive in their normal style.

Novice drivers had on average shorter fixation durations, except on the suburban drive, and made fewer fixations than experienced drivers. Additionally, the results of the fixation position analysis showed that experienced drivers increased their search in the horizontal axis relative to the rural route on the dual carriageway, and to a lesser extent on the suburban route. The novice drivers tended to maintain the same level of horizontal search throughout all the road types, at a similar level of horizontal search produced by experienced drivers on the suburban route. Finally, experienced drivers exhibited less vertical spread of fixations on the dual-carriageway and the rural route compared to the novice drivers.

These results taken together suggest that experienced drivers select visual search strategies according to the complexity of the roadway, whereas the visual search strategies of novice drivers change little from roadway to roadway, indicating that the strategies of novice drivers are inflexible to changing visual and cognitive demands. Specifically, longer fixation durations are often taken to indicate increased processing time due to amongst other things the complexity of the visual scene (Rayner, 2009). The novice drivers' longer fixation durations on the more demanding dual-carriageway reflect the increased cognitive demand; whilst the experienced drivers' shorter fixation durations and increased number of fixations show that they are using a strategy, likely developed though experience, to sample more of the scene in order to fixate on areas where they can get more task relevant information. In addition, on the more demanding roadway the experienced drivers expanded their search strategy along the horizontal axis and to a smaller extent the vertical, whilst the novice drivers did not vary the size of their visual searches along either the horizontal or vertical axis across the three routes, even as the cognitive load of the task increased. This can be interpreted as the adoption of an inflexible visual search strategy. Finally, the increased fixation durations of the novice drivers could be a result of perceptual narrowing, which has been shown to affect visual attention (Henderson, Pollatsek, & Rayner, 1989; Henderson, 2003), such that because the dual carriageway is highly demanding then the effective peripheral field of vision may shrink.

Overall, the implications of this study are that, in line with the earlier research, novice drivers show poorer visual attention and visual search strategies than experienced drivers. There is an increased inflexibility to a novice driver's search strategy across road type and as cognitive load increases. However, a simple note of caution should be exercised when interpreting the results because the cognitive load in this task was not clearly controlled so there could have been other factors that influenced the load of the task. For example, it is not known if drivers were familiar with the route, had similar weather conditions or traffic at the time of day when the study took place, and if they were matched across groups. These are all factors that could influence how demanding the task would be to complete. Additionally, there is no objective or subjective measure of the actual task demands, as this was not recorded or reported in the study methodology. It is possible that the task demand was perceived to be similar by the two groups and that another factor not measured here influenced the results of this study.

Indeed, interestingly, somewhat contrasting results emerged from a driving simulator-based study which allowed more careful control of many of these factors. Konstantopoulos, Chapman, and Crundall (2010) examined the visual search behaviour and search strategies of novice and experienced driving instructors to determine the specific strategies each group uses when they drive. In this study, they recorded the eye movements of driving instructors and learner drivers while they drove three virtual routes that included day, night and rain routes. Ten driving instructors (mean driving experience: 34 years) and 11 novice drivers (mean number of driving lesson hours: 24 hours) drove three predetermined routes in a driving simulator whilst having their eyes tracked. The key eye tracking measures were number of fixations, mean fixation duration, standard deviations of the horizontal and vertical fixation locations, and area of interest (AOI) fixations (rear, side mirrors and speedometer). The driving routes were geographically the same and the only difference between routes was visibility with the engagement of three conditions: day, night and rain. Auditory directions were presented to guide participants along the route, as well as arrows at the bottom of the screen. The results showed that driving instructors made more fixations than novice drivers on average (mean = 634 vs. 519 fixations), and this did not vary between the three road conditions. The fixation duration of driving instructors was shorter than novice drivers (431 vs. 567 ms) and this did not vary between the three road conditions. Next the horizontal spread of fixations was broader for the driving instructors than the novice drivers (10.6^{o} vs. 6.2^{o}) and again there was no effect of road condition. This result contrasts with the previously discussed findings (e.g. Crundall & Underwood, 1998) showing adaptation of experienced drivers' scan patterns according to the conditions. This result potentially highlights one of the difficulties in comparing the results from simulated and real world driving studies. Next, the vertical spread of fixations did not vary between group or road condition. The AOI analysis showed that driving instructors made more fixations to the right mirror (11.9 vs 2.1) but not the left than novice drivers. There was no difference in rear mirror fixations between the groups, except in rain and night conditions, where driving instructors made more fixations to the rear-view mirror than novices. Finally, the driving instructors made fewer fixations on the speedometer than novice drivers (3.3 vs 13.7).

Taken together these results indicate a similar pattern of results to the studies discussed above. Experienced drivers tend to make more fixations, have shorter fixation durations and exhibit a wider horizontal spread of fixations than novice drivers, which suggests that experienced drivers are better equipped to search for and process task relevant driving information than novice drivers. Novice drivers – likely due to their lack of driving experience and knowledge – are inefficient in their search patterns and take longer to process the same visual information as experienced drivers.

Young Novice Drivers and Dual Task Behaviour

In this next section I will discuss some of the influence that dual task workload can have on driving performance in young novice drivers. In a simulator study, Cantin, Lavallière, Simoneau, and Teasdale (2009) examined if the mental workload of young and experienced drivers varies with the difficulty of the driving scenario. Ten younger drivers (mean age: 24 years) and 10 experienced drivers (mean age: 65 years) drove through a simulated driving environment continuously for 26 km, which included rural and urban scenes. A probe reaction time (RT) technique was used to measure the workload while driving. Participants were instructed that the primary task was driving and the secondary task was to verbally respond every time a sound was heard. The secondary task RT probes were given in a baseline static condition and in three different driving scenarios: driving on straight roads at a constant speed; approaching intersections for which the driver had to stop the car; and overtaking a slower vehicle. RTs in the baseline condition were not significantly different between the groups (roughly 400ms for both). Both groups exhibited longer RTs as the complexity of the driving scenario increased (i.e. RTs in the region of 600-1000ms). Interestingly though, the experienced driver group exhibited longer RTs than the younger drivers for probes occurring during the overtaking manoeuvre (700 vs 1100ms). This suggests that the overtaking manoeuvres led to a greater mental workload for experienced drivers than for younger drivers and that more complex driving contexts required more of the available cognitive resources for experienced than younger drivers. However, the results of this study should be interpreted with caution as the relatively aged experienced drivers used in Cantin et al.'s study (mean age: 65 years) may not be representative of the experienced driving population and it is likely that a diminishing cognitive capacity in these experienced drivers could account for the findings. This illustrates the importance of comparing young novice drivers with experienced drivers who are also relatively young, and this is one of the aims of the current work.

Young Drivers' Executive Function

Executive control functions (such as response inhibition, working memory, and mental set shifting) are key to most everyday tasks (Diamond, 2013; Monsell, Stephen and Driver, 2000; Verhaeghen & Cerella, 2002). Executive control is therefore an important factor in the many tasks that are required for safe driving (e.g. navigation, speed control, lane changing, hazard perception). Given the elevated crash risk posed by novice drivers, research has examined whether young drivers might have executive control deficits relative to older drivers, which could impact on their driving ability. For example, Mäntylä, Karlsson, and Marklund (2009) examined whether individual differences in executive function (EF) might have selective effects on driving performance, such that particular aspects of EF might contribute more to novices' driving errors than other aspects of EF. Fifty participants (mean age: 17 yrs. old) completed six EF tasks (e.g. Stroop and flanker tasks) whilst also driving a simulated vehicle on a computer screen. The driving task, lane control task (LCT, involved staying in a lane as accurately as possible and occasionally changing lane as indicated by the traffic signs appearing on the computer screen. The speed of the vehicle was automatically controlled and each trial took 10 minutes. The LCT score was not correlated with the inhibition and set shifting components of EF. However, the LCT score did correlate with the updating component of EF (which reflects the ability to update mental representations of a task), such that participants with low performance in the updating component made greater errors in the simulated driving task. In summary, these results show that young drivers with low EF, specifically in mental updating, could be at greater risk for driving accidents than young drivers with better mental updating abilities. However, because this study did not include comparison groups of older and/or young experienced drivers, it is unclear whether people of any age who score low on the updating component of EF make more driving errors, or whether this pattern only occurs in young novice drivers. In addition, once again, the results of this study should be interpreted with caution as the driving task was not very similar to any of the tasks that would be required to drive a vehicle on the road or even in a high fidelity driving simulator.

Young Novice Drivers' Hazard Perception

It seems likely that the differences in patterns of visual attention exhibited by novice drivers compared to more experienced drivers described above might result in different levels of awareness of the driving context and its ongoing development. Indeed, there is substantial evidence to suggest that novice drivers are worse at hazard perception than more experienced drivers. Novice drivers are more likely than older drivers to underestimate dangerous situations or not be able to recognize hazardous situations (Jonah & Dawson, 1987). Additionally, novice drivers are also more likely than adults to make critical decision errors that lead to serious crashes (McDonald, Curry, Kandadai, Sommers, & Winston, 2014).

To examine what specific factors influence hazard perception in young novice drivers Scialfa et al. (2011, 2012) conducted two studies. In the 2011 study, Scialfa and colleagues showed novice and experienced drivers a series of short video scenes and asked them to indicate the presence of a traffic conflict that would lead to a collision between the "camera" vehicle and another road user. 29 novice drivers (less than 6 months driving experience) and 146 experienced drivers (more than 2 years driving experience) viewed 95 driving scenes through several different environments (e.g. urban, suburban, motorway etc.) of which 64 had a driving hazard, and were asked to respond as soon as they detected the hazard. Novice (N) drivers missed more hazards (8.6%) than experienced (E) drivers (5.2%), and importantly had slower hazard perception RTs (3160ms) than experienced drivers (2760ms). The groups did not differ in their false alarm rate (N = 3.6% v. E = 4.5%). These effects of driving experience were independent of age and the group differences in hazard perception RT occurred in addition to differences in general speed of responding. In summary, the results of this study support earlier research in demonstrating that novice drivers are slower to recognise hazards but do not differ in their false alarm rates.

In their 2012 study, Scialfa and colleagues used picture images of potentially hazardous driving scenes to examine the effect of driving experience on hazard perception. 29 novice drivers (less than 6 months driving experience) and 27 experienced drivers (more than 2 years driving experience) viewed 120 images in which 100 had potential traffic conflicts (defined as situations in which a collision (or near collision) between the driver and another road user would occur, or had the potential to occur, unless the driver took evasive action such as slowing, stopping, or steering). Participants were asked to rate on a 5 point Likert scale ‘How critical is the hazard to safe driving?’ where 1 represented ‘no hazard’ and 5 represented an ‘extremely critical hazard’. Participants were asked to make these ratings as quickly as possible so that the researchers could compute a hazard perception RT. The results showed that novice drivers were significantly slower than experienced drivers in categorising an image as hazardous (2060 ms vs 1750 ms), novice drivers missed more hazards than experienced drivers (89.9% vs. 95.8% correctly detected), and categorised the 100 hazardous scenes on average as less hazardous than experienced drivers (2.6 vs. 2.9 Likert score). In summary, these results show three important factors that affect novice drivers’ hazard perception ability. One, novice drivers are slower to respond to driving hazards than experienced drivers, likely because of a lack of experience of driving hazards. Two, they are more prone to missing a driving hazard than experienced drivers, again, likely because of a lack of exposure to a variety of driving hazards. Three, novice drivers categorise driving hazards as being less serious than experienced drivers do, which may place them in riskier situations than experienced drivers.

Similar results emerged from a hazard perception study that included eye tracking. Borowsky and Oron-Gilad (2013) examined the effects of driving experience on hazard awareness and risk perception skills in novice, experienced and very experienced drivers whilst they had their eyes tracked. 27 novice drivers (M = 1.5 months' driving experience), 30 experienced drivers (M = 7.6 years' experience) and 25 very experienced drivers (M = 23.5 years' driving experience) performed three consecutive tasks. First, they observed 10 short movies of real-world driving situations and responded with a button press each time they identified a hazardous situation. Second, they observed one of three possible sub-sets of 8 movies (out of the 10 they had seen earlier) and were asked to categorise them according to the similarity of the hazardous situation (whether the hazards were of a similar severity). Third, they observed the same sub-set for a third time and following each movie were asked to rate its level of hazardousness and the likelihood of there being a crash. The results showed that very experienced drivers were more likely than young novice drivers to report an intersection as hazardous, whereas experienced drivers were somewhere in between and did not differ with the other two groups. The very experienced drivers were also better at responding and fixating on the hazard in the more hazardous scenarios (e.g. busy junctions) than the novice drivers. Finally, the results showed that novice drivers underestimated the likelihood of a crash and the severity of its outcome more than both the experienced and very experienced driver groups. Specifically, they judged the severity of a crash's outcome as less risky than the likelihood of the outcome, which contrasts with the experienced and very experienced drivers who prioritised the severity of the outcome as being of higher risk than the likelihood of a crash. In summary, novice drivers exhibit a less developed understanding of crash severity and likelihood than experienced and very experienced drivers. Additionally, they were also poorer at identifying the areas of a scene upon which they needed to fixate in order to gather the information needed to make a safe manoeuvre.

One of the key questions about young novice drivers' hazard perception is how it might improve over time. Sagberg and Bjørnskau (2006) conducted a study to investigate driving risk decreases in relation to improved hazard perception skills and gender differences. They used a video based hazard perception test to measure reaction times to hazards appearing in 31 traffic scenes in three young novice driving groups (who had held a licence for 1, 5, and 9 months) and a group of drivers who had held their licences for several years. The participants were asked to respond to whenever a hazard occurred. A hazard was defined "as any motion by some other road user, which could possibly develop into a hazard, and for which the driver had to be especially prepared for taking some evasive action in terms of braking or steering" (Sagberg & Bjørnskau, 2006, pg 2). The results showed that the least experienced drivers (1 month) took the longest to respond to the driving hazards, compared with the 5 and 9 month groups. In further analysis, it was found that more experienced drivers (5 and 9 month) did not differ significantly in mean reaction times for all driving situations. However, female drivers had significantly shorter RTs to just under 50% of the individual situations (more complex driving scenarios), and the shorter RT difference was found in female novice drivers in comparison to male novice drivers. A clear reason for the lack of difference as a result of experience is that the 5 and 9 month group were not significantly more experienced and that the differences between the groups was not large because of this fact. A better contrast would be to compare between driver groups with a wider range of experience, as I did in this study and in the next chapter (Chapter 5). The results of the studies cited above indicate three issues. First, hazard perception RT performance may differ between novice and experienced drivers. Second, a small gender difference can be observed in certain more complex driving situations. Third, hazard perception is potentially a minor factor in explaining the risk decrease among novice drivers due to experience and other factors which may play a more important part in young driver crash risk. These results taken together may indicate only that more complex driving hazards cause poorer hazard perception in novice drivers and that sex differences may exist in the novice driving population, placing male drivers at higher risk of crashing than females. In fact, data shows that male young drivers have more crashes than female young drivers (Department for Transport, 2015, 2016).

The use of hazard perception training interventions with young inexperienced drivers can improve hazard perception (e.g. Chapman, Underwood, & Roberts, 2002; Pollatsek, Narayanaan, Pradhan, & Fisher, 2006). For example, McKenna, Horswill and Alexander (2006), found across 3 studies that a brief 20-minute hazard perception skill-based training program using video-based driving simulations could improve hazard perception skill in inexperienced drivers and reduce their risk taking. Using a video based simulator the hazard perception ability of young novice drivers was measured before and after a skill-based training intervention. 91 novice drivers (mean age = 19), with less than 3 years of driving experience after passing their test (M = 1.5 yrs.) were randomly assigned to one of two groups (trained and un-trained) and completed a number of driving behaviour questionnaires and video based hazard perception tests. The trained group watched several video-based hazard perception training videos, with commentary referring to potential hazards throughout the video. The un-trained group did not see these videos. The results showed that the trained group reported being significantly less likely to speed or take risks (measured through questionnaire responses and responses to videos seen by both groups) than the un-trained group. Additionally, the mean self-perceived skill (0-10) for the trained group (6.52) was not significantly different from the mean rating of the untrained group (6.76) indicating that although the trained group did not perceive themselves to have improved their hazard perception skill, the training intervention had improved their abilities. In the final part of the study, to examine if the same pattern of results emerge when using a more naturalistic training regime, the hazard perception abilities of experienced police drivers was studied. Police driver training operates at three levels: basic, standard, and advanced. The police drivers completed a similar task as the novice drivers had above. The results showed a similar pattern as in the novice driver study, such that police drivers with advanced training performed better in hazard perception tests than those with basic or standard level training. In summary, these results show that training interventions for both novice and experienced drivers can improve hazard perception ability. However, what is not clear is the longevity of the effect of these training interventions, although the results of the advanced police drivers hint towards a long-term effect of training interventions.

Similar effects were observed in a more recent study by Isler, Starkey and Sheppard (2011). 36 young drivers (16-18 years) were allocated to one of three groups for five days of training. One group received higher-order driving training, to improve perceptual, motivational, risk taking and insight skills. The second group received basic driving handling training, and a third control group received no training. All participants had their driving skill measured by a battery of tests, such as hazard perception and risk taking, before and after training. The group that received higher-order training improved across all aspects measured. In particular, a statistically significant improvement in relation to visual search and hazard perception was identified. The group that received vehicle training only showed improvements in vehicle and speed control but not hazard perception, compared to the control group. A note of caution should be taken, in that although the novice drivers in this study exhibited overall an improvement in hazard perception, the lack of comparison with another driving group makes it difficult to know if these effects are group or population specific. In sum, hazard perception training for young and inexperienced drivers can improve their overall hazard perception abilities.

Predictions

The evidence presented above paints a clear picture of the type of behaviour novice drivers exhibit whilst completing real world, simulated or lab based driving tasks. Specifically, novice drivers exhibit longer RTs to detect hazards, miss more hazards and miscategorise the severity of driving hazard for a variety of factors (e.g. lack of experience of driving hazards, poor dual task performance compared to young experienced but not older drivers, and poor hazard perception strategies). In the present study, I examined young novice drivers' identification of driving hazards under different levels of cognitive load. Based on the research described above, I predict that as the cognitive load increases, RTs to driving hazards will increase.

Method

Participants

Thirty young novice drivers were recruited for the study through Royal Holloway (mean age = 19, SD = .9, 27 females). The sample size was derived from prior research, constituting a typical size for studies in this area (e.g. Galpin, Underwood, & Crundall, 2009). 28 participants were right-handed. All had a full UK driving licence and a mean driving experience in months of 6.9 (SD = 3.6). The average miles driven per year was 2008.3 (SD = 2094). Participants were naïve to the purpose of the experiment, and had not previously taken part in driving experiments at RHUL. All participants had normal or corrected-to- normal visual acuity, gave informed consent before participating, and were compensated with £10 for participation in the study. All procedures were reviewed and approved by the Departmental Ethics Committee.

Stimuli and Design

The apparatus, stimuli, design and procedure were all exactly as described in Chapter 2.

Results

In all subsequent *F*-tests, the Greenhouse-Geisser correction was applied if Mauchly's test of Sphericity was significant, and the Benjamini-Hochberg false discovery rate procedure was used for all post-hoc tests. The false discovery rate was set at 0.1 for all multiple comparisons (Benjamini & Hochberg, 1995).

Data Preparation

Workload measures.

As in previous chapters, number probe responses were calculated from the first four responses per trial for each participant. All responses with empty cells were removed (3.2%), and all incorrect responses were removed (6.4%) for the RT analyses. The final dataset consisted of 4569 correct number probe responses (low load = 2290; high load = 2279), and overall 9.6% of the data were removed for the RT analyses.

Hazard detection measures.

Also as in previous chapters, in order to ensure that the hazards were clearly detectable across all the videos for all the participants who took part in this study, data from videos with overall hazard response accuracies of less than 70% were removed from the analysis (14 videos, 20.7% of the data). Subsequently, response accuracy across all load conditions was high (M = 82%, SD =.39). All inaccurate responses were removed for the RT analyses (14.6%). In order to identify the earliest point at which a hazard could be detected, the fastest correct response was identified for each video and this time point was made the baseline against which all other hazard responses were measured (in 9 videos, the fastest response came from the participants; in the remaining videos, the fastest response came from an evaluator or previous experiment benchmark). The final dataset consisted of 1164 correct hazard responses, and overall 35.3% of the data were removed for the RT analysis.

Data Analysis

Workload measures.

In order to examine the effectiveness of the load manipulation, a repeated measures ANOVA was conducted on the number probe RTs, using the within-subjects factor of load condition (low vs. high). There was a significant effect, $F(1, 29) = 315.4$, $p < .001$, $\eta_p^2 = .92$, indicating that RTs under low load (M = 571 ms, 95% CI = 538 – 605) were significantly faster than under high load (M = 959 ms, 95% CI = 907 – 1010), which confirms that the load manipulation was successful. However, the accuracy analysis revealed no significant performance difference under low load (M = 95.5%, 95% CI = 94 – 97) versus high load (M = 93.2%, 95% CI = 91 – 95.5), $F(1, 29) = 3.6$, $p = .067$, $\eta_p^2 = .11$.

The RSME subjective measure of workload was also tested using a repeated measures ANOVA with the within-subjects factor of load condition (no load, low load and high load). There was a significant effect of load condition, $F(2, 58) = 26.9$, $p < .001$, $\eta_p^2 = .48$ and post-hoc tests confirmed that each condition was different from the others (all $p < .001$; see figure 4.1). These results indicate that the workload manipulation also worked as predicted from a subjective point of view.

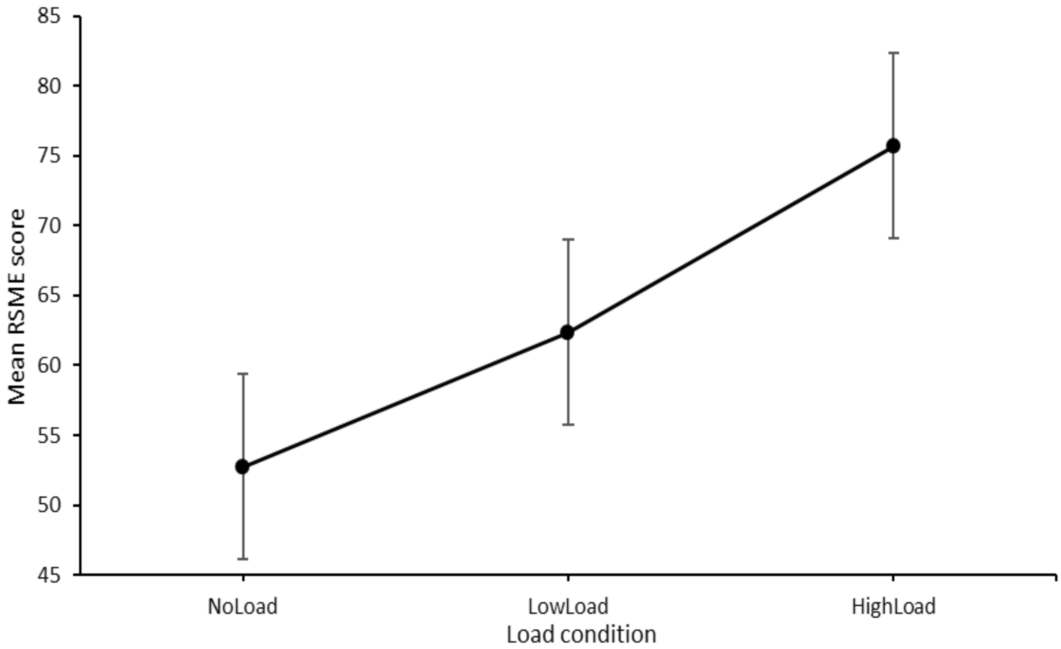

Figure 4.1. Mean RSME scores (± 95% CI) as a function of load condition.

Hazard detection measures.

Two repeated measures ANOVAs were conducted on hazard detection RTs and accuracy using the within-subjects factor of load condition (no load, low load and high load). There was no significant effect of load condition on mean hazard detection RTs, $F(2, 58) = 3.48$, $p = .45$, $\eta_p^2 = .03$ (see Figure 4.2). Accuracy across the three load conditions was high (no load - M = 83.6%, 95% CI = 76.8 - 90.3; low load - M 79.1%, 95% CI = 72.3 - 85.8; high load - M = 81.1%, 95% CI = 74.7 – 87.6), and did not vary significantly across load conditions, $F(2, 58) = 1.5$, $p = .28$, $\eta_p^2 = .05$.

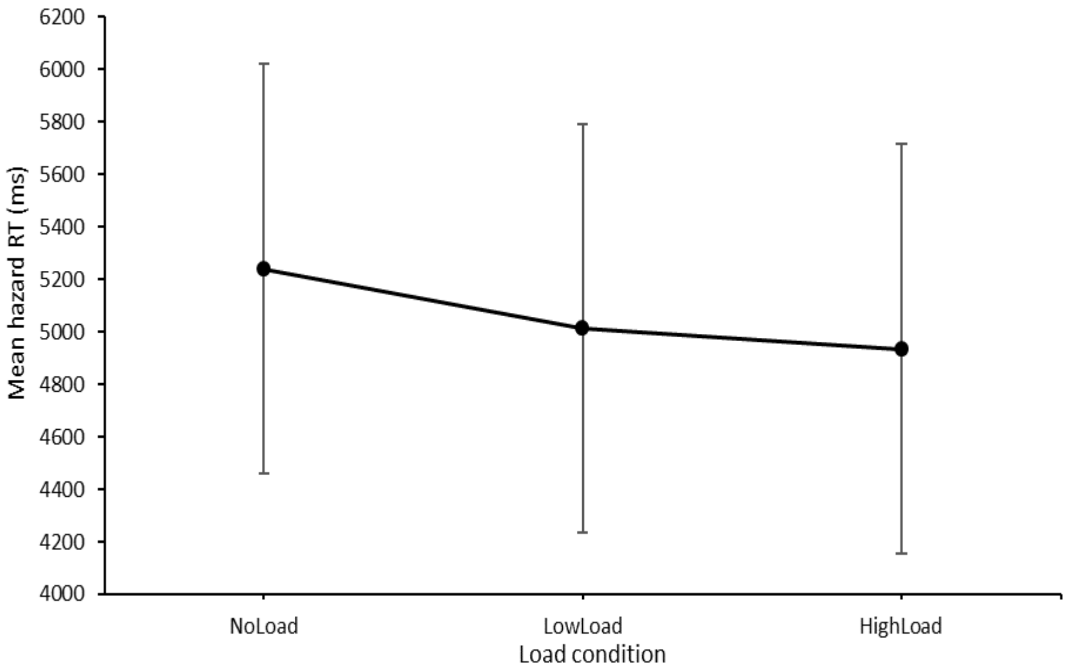

Figure 4.2. Mean correct response times (± 95% CI) in milliseconds as a function of load condition.

In the final analysis participants' RT and accuracy data for hazard detection were then converted to inverse efficiency (IE) scores in order to correct for any potential speed–accuracy trade-offs that might have been present. A repeated measures ANOVA was conducted on IE scores using the within-subjects factor of load condition (no load, low load and high load). The analysis revealed no significant main effect of load condition. Participants' hazard detection performance was not significantly different between no load (M = 8210, 95% CI = 4280 - 12140), low load (M = 7385, 95% CI = 4878 - 9892) and high load conditions (M = 7044, 95% CI = 4592 - 9496), $F(1.3, 38.8) = 1.32$, p =.27, η_p^2 =.04.

Discussion

The results of the present study demonstrate that novice drivers' RTs for detecting driving hazards in realistic video stimuli did not change significantly as dual task workload increased. Additionally, there was no significant change in hazard detection accuracy as workload increased. These finding are in not in line with previous research on dual task interference which typically find significant costs in dual (vs. single) task conditions. Taken in total the results of this study show that even as secondary task workload increases there is no detrimental impact to primary task performance. This indicates that young novice drivers might be partially better protected from dual task interference because of their younger age and they might be able to switch between secondary task of varying workload with little drop in their primary task performance. These questions will be considered in more detail in Chapter 5, which compares performance between the different driver groups tested in Chapters 2, 3 and 4.

It is possible that the lack of dual task effects seen here might stem from the participants having relatively high working memory capacity (WMC). For example, Wood, Hartley, Furley and Wilson (2016), examined the influence of individual differences in WMC on hazard perception performance in a simulated driving task. They found that WMC scores significantly predicted hazard perception performance in the dual task condition (R^2 =.11) even when controlling for hazard perception performance. Participants with High-WMC performed considerably better than the Low-WMC group in the dual task condition, such that, while the high-WMC group maintained their performance across both conditions, the low-WMC group performed worse under dual task compared to control conditions. Thus, one possible explanation for the lack of dual task effects in the present study is that the participants sampled may all have had reasonably high WMC (this seems plausible, for example, given that they were all recruited from a university student sample).

Chapter 5 – Differences Between Younger Experienced, Older and Novice Drivers in the Effects of Workload on Hazard Detection Performance

Introduction

Because experiments in Chapters 1-3 were run and analysed one after another, the data analysis was slightly different for each group tested. For example, RTs were calculated with respect to the earliest point at which any participant correctly identified the hazard, and this point could change across the experiments. For this reason, in order to make meaningful comparisons between the different groups of drivers, it is important to process all the data in exactly the same way. This is the purpose of the current chapter: by pooling the data from across all three experiments reported so far, I will be able to examine the differences in hazard response as a function of secondary task load between the different driver groups (younger experienced, novice and older drivers). The main hypotheses for this chapter are that there will be significant differences in primary task RT and task accuracy between the younger experienced, novice and older drivers as the task workload increases. Specifically, it is predicted that the younger experienced drivers will exhibit shorter RTs and better accuracy in the primary and secondary task than both novice and older drivers. Additionally, separate comparisons between younger experienced and novice drivers, and younger experienced and older drivers, might show shorter RTs and better accuracy in the primary task as secondary task load increases for the younger experienced drivers.

Method

The methods were identical for all three groups, as described in Chapter 2. The data processing also proceeded in the same way as for the previous experiments, but the raw data from all participants, trials and experiments were now pooled together (5400 trials). Any videos with 70% or less accuracy were removed (21 videos, 34% of the data), as previously described in Chapter 2. The RTs were benchmarked so that the fastest correct hazard detection across all three groups for each video was used for the analysis. For the RT analysis all inaccurate responses were removed (7.9%), leaving 3138 trials. Therefore, in total 41.9% of the data was removed.

Results and Preliminary Discussion

In all subsequent *F*-tests, the Greenhouse-Geisser correction was applied if Mauchly's test of Sphericity was significant, and I used the Benjamini-Hochberg false discovery rate procedure for all post-hoc tests. The false discovery rate was set at 0.1 for all multiple comparisons (Benjamini & Hochberg, 1995).

Workload Measures

Rating scale mental effort (RSME).

To confirm the effectiveness of the load manipulation, and compare how this was experienced between the different driver groups, a mixed ANOVA was conducted on the RSME data, using the within-subjects factor of load condition (no load, low load, high load) and experimental group as the between-subjects factor (young experienced, novice and older population). As expected, given the significant effects of load seen throughout this thesis so far, there was a significant main effect of load condition, $F(1.6, 139.4) = 91.3$, $p < .001$, $\eta_p^2 = .51$. All post-hoc comparisons between the different load conditions were significant (all $p < .001$), indicating that perceived workload increased reliably from the no load condition to the low load condition and again to the high load condition (figure 5.1). There was no significant effect of experimental group, $F(2, 87) = 2.07$, $p = .13$, $\eta_p^2 = .04$, indicating that the groups experienced a similar overall level of subjective workload. However, there was a significant interaction between load condition and experimental group, $F(3.2, 139.4) = 5.04$, $p = .002$, $\eta_p^2 = .10$. Post-hoc comparisons showed a significant difference in RSME score in the no load condition between young experienced (score = 42.8) and older (score = 56.5) drivers ($p = .01$), all other comparisons were non-significant ($p > .05$). Thus, whereas the dual task conditions (low load and high load) were rated as similarly demanding by all groups, older drivers rated the single task condition (no load) as significantly more demanding than did the younger experienced drivers. This might reflect an overall reduced capacity in the older drivers, such that the single task is already experienced as relatively demanding.

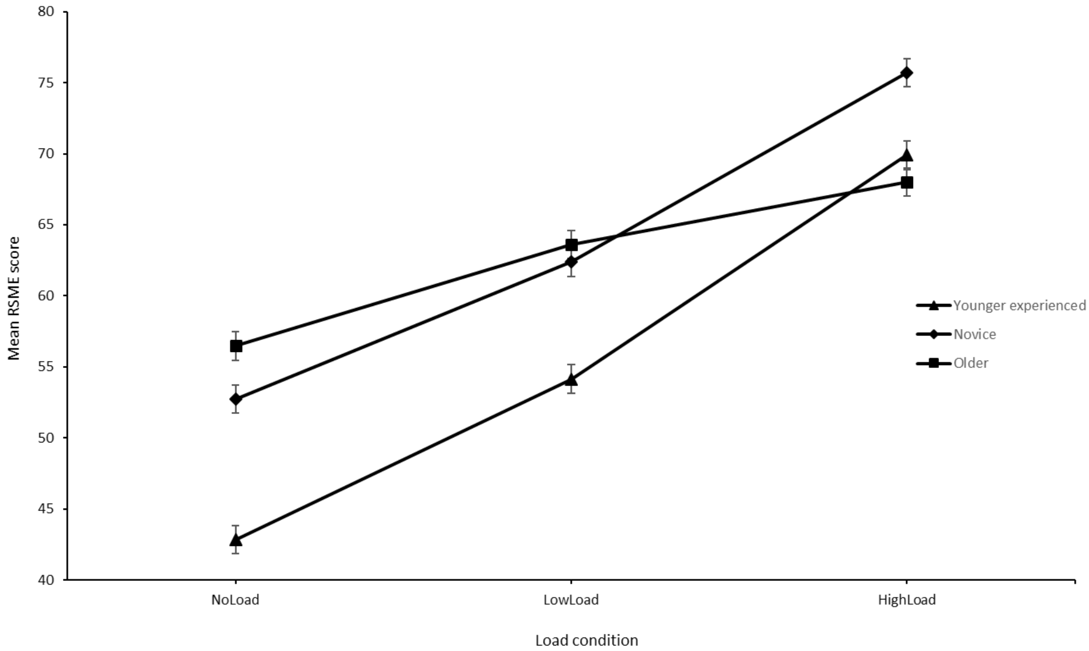

Figure 5.1. Mean RSME scores (± 95% CI) as a function of load condition and driver group.

Number probe accuracy.

To compare responding on the number probe task across the different driver groups, a mixed ANOVA was conducted on the number probe accuracy, using the within-subjects factor of load condition (low vs. high) and experimental group as the between-subjects factor (young experienced, novice and older population). As would be expected given the results of the previous chapters, the effect of load condition was significant, $F(1, 87) = 19.7$, $p < .001$, $\eta_p^2 = .18$, with higher accuracy under low load (93.9%) than under high load (90.5%). The main effect of group was also significant, $F(2, 87) = 25.5$, $p < .001$, $\eta_p^2 = .37$, with significant differences between young experienced (97.8%) vs. older drivers (84.4%) and novice (94.4%) vs. older drivers (84.4%) (both $p < .001$). There was no difference between young experienced and novice drivers' performance, $p > .05$ (figure 5.2). Thus, perhaps unsurprisingly, older drivers were less accurate at responding to the number probe than both young experienced and novice drivers. It will be important to bear this in mind when considering performance on the hazard detection task, as it is possible that older drivers sacrificed performance on the number probe task in order to achieve good performance on the hazard detection task. The interaction between load condition and experimental group was not significant, *F(2, 87)* = .75, *p* =.48, η_p^2 =.02.

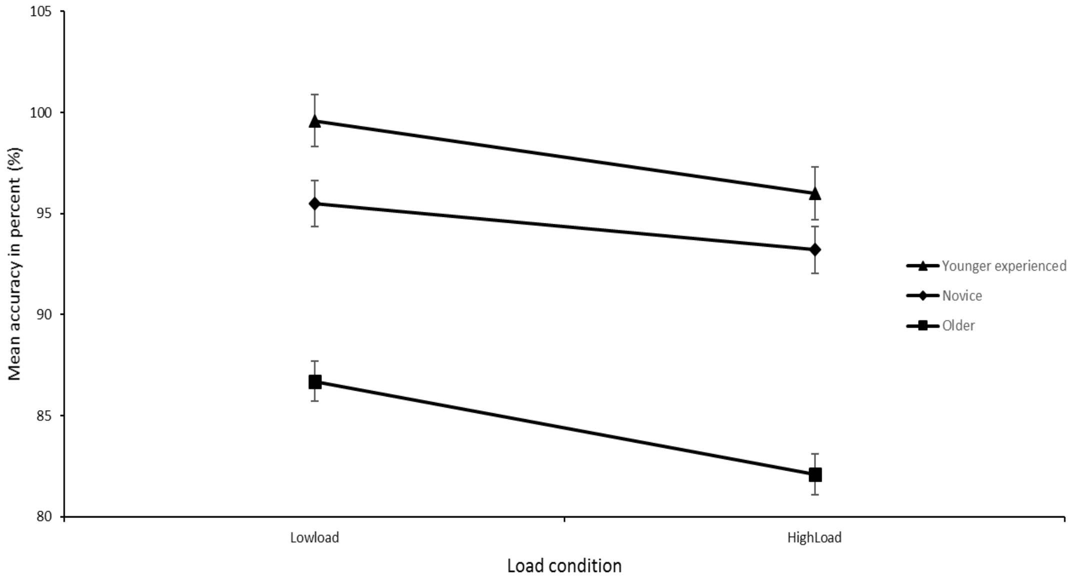

Figure 5.2. Mean number probe accuracy (±95% CI) as a function of load condition and driver group.

Number probe response time.

To compare the speed of responding to the number probe between the different groups, a mixed ANOVA was conducted on the probe RTs, using the within-subjects factor of load condition (low vs. high) and experimental group as the between-subjects factor (young experienced, novice and older population). In line with the accuracy results and with the results from the previous chapters, the main effect of load condition was significant, *F(1, 87)* = 1040.4, *p*< .001, η_p^2 =.92, with faster responding under low load (613 ms) than under high load (1004 ms). The main effect of group was also significant, *F(2, 87)* = 11.8, *p* <.001, η_p^2 =.21. In line with the number probe accuracy results, there were significant differences between young experienced (773 ms) and older drivers (888 ms) and between novice (765ms) and older drivers (888ms, all *p*<.001) but not between the young experienced and novice driver groups *p*>.05. Thus, once again, when comparing the older group's hazard detection performance with that of the two other groups, it will be important to note that the older group displayed significantly slower responses on the number probe task, as well as reduced accuracy, and thus may have been sacrificing performance on this task in order to maintain hazard detection performance. The interaction between load condition and group was not significant, *F(2, 87) =* .60, *p* =.55, η_p^2 =.01 (figure 5.3).

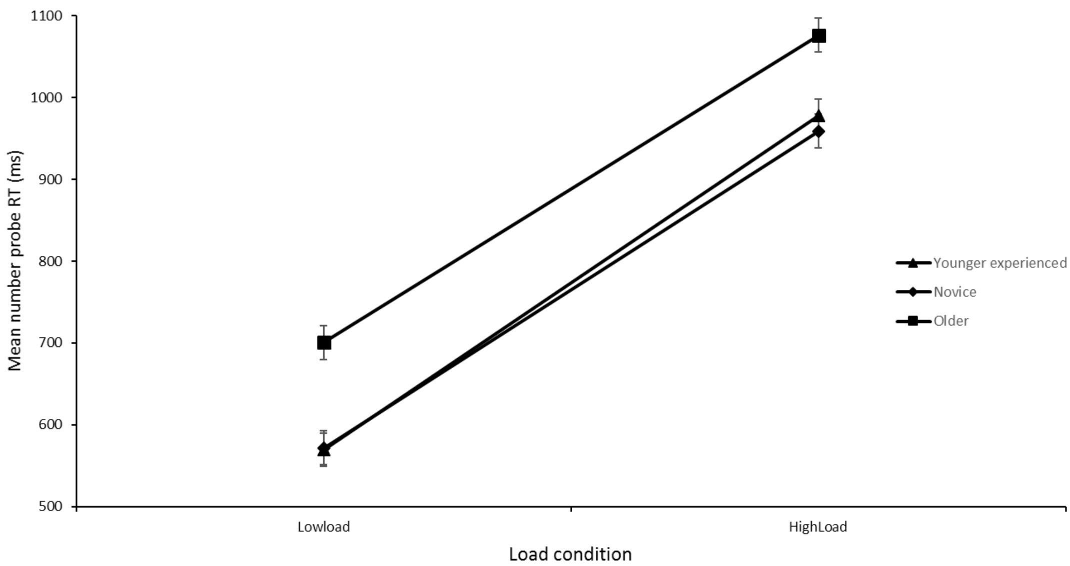

Figure 5.3. Mean number probe RT (±95% CI) as a function of load condition and driver group.

Hazard Detection Measures

For the hazard detection analyses, identical mixed ANOVAs were conducted on the RTs, accuracy and inverse efficiency data using the within-subjects factor of load condition (no load, low load and high load), and experimental group as the between-subjects factor (young experienced, novice and older population).

Hazard detection RT.

There were no significant main effects of load condition, *F(2, 174)* = .94, *p* =.39, η_p^2 =.01 or of group , *F(2, 87)* = 2.19, *p* =.12, η_p^2 =.05. However, the interaction between load condition and group was significant, *F(4, 174)* = 3.19, *p* =.015, η_p^2 =.07. Post hoc comparisons identified a significant difference in RT in the no load condition between novice (4893 ms) and older (3583 ms) drivers (*p* =.04). No other comparisons were significant (*p*>.05) (figure 5.4). Planned comparisons between young experienced and novice drivers at each load condition showed that there was a significant difference in RT between the driver groups during the low load condition (*p* =.03) but not during the no or high load conditions (*ps* >.07). Additional planned comparisons between young experienced and older drivers at each load condition found no significant differences in RT at any of the load conditions (*ps* >.06).

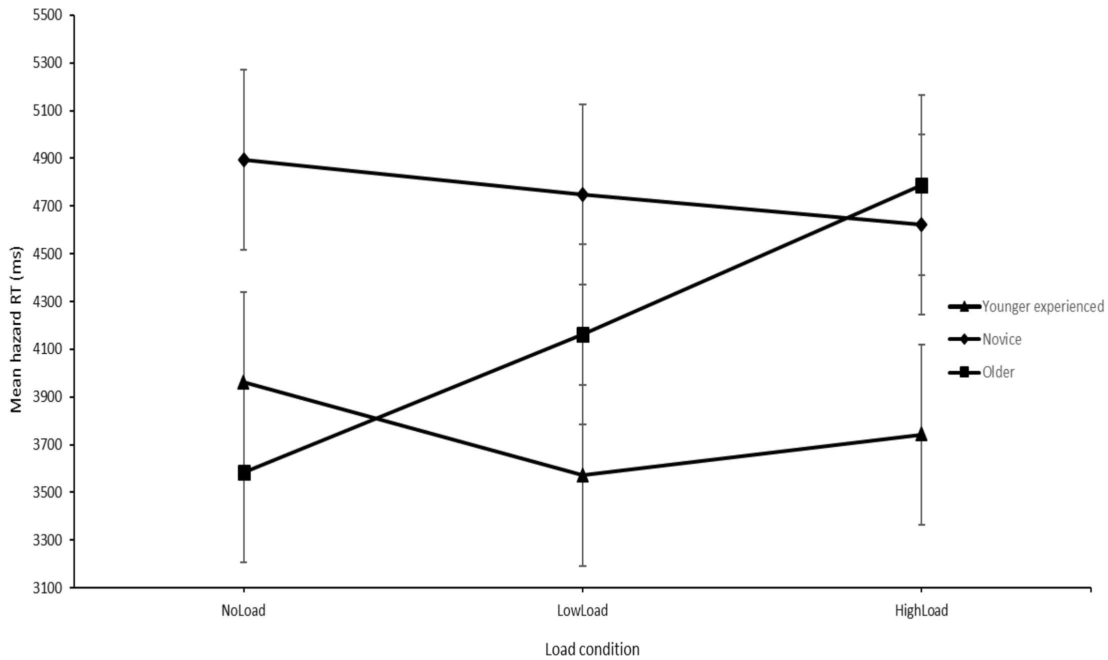

Figure 5.4. Mean hazard detection RT (± 95% CI) as a function of load condition and driver group.

Hazard detection accuracy.

The main effect of group was significant, *F(2, 87)* = 5.62, *p* =.005, η_p^2 =.11. Post hoc comparisons identified significant differences between young experienced (93.5%) and older (82.6%) drivers and between young experienced (93.5%) and novice drivers (84.8%, all *p*<.05). Thus, interestingly, both the novice drivers and the older drivers were significantly less accurate at hazard detection than the young experienced drivers, regardless of the level of load. There was also a significant effect of load condition, *F(2, 174)* = 3.9, *p* =.02, η_p^2 =.04. Post hoc comparisons identified significantly higher accuracy in the no load condition (87.9%) than in the low load condition (84.3%, *p* =.02). This is in line with the wealth of literature that has demonstrated performance costs under dual (vs. single) task conditions.

However, in contrast to this literature, accuracy was worse with a low load in the number probe task (84.3%) than it was with a high load in that task (88.7%, *p* =.02). This improved hazard detection accuracy under a high load in the additional number probe task mirrors the result from Chapter 2 and perhaps suggests that the added engagement provided by the high load number probe task also boosts responding in the concurrent hazard detection task. The interaction was not significant, *F(4, 174)* = 1.7, *p* =.15, η_p^2 =.04 (figure 5.5). Planned comparisons between young experienced and novice drivers at each load condition showed that there was a significant difference in accuracy between the driver groups during the low load condition (*p* =.004) but not during the no or high load conditions (*ps* >.06). Additional planned comparisons between young experienced and older drivers at each load condition revealed a significant difference in accuracy under no load (*p* =.001) and low load (*p* <.001) but not high load (*p*>.10).

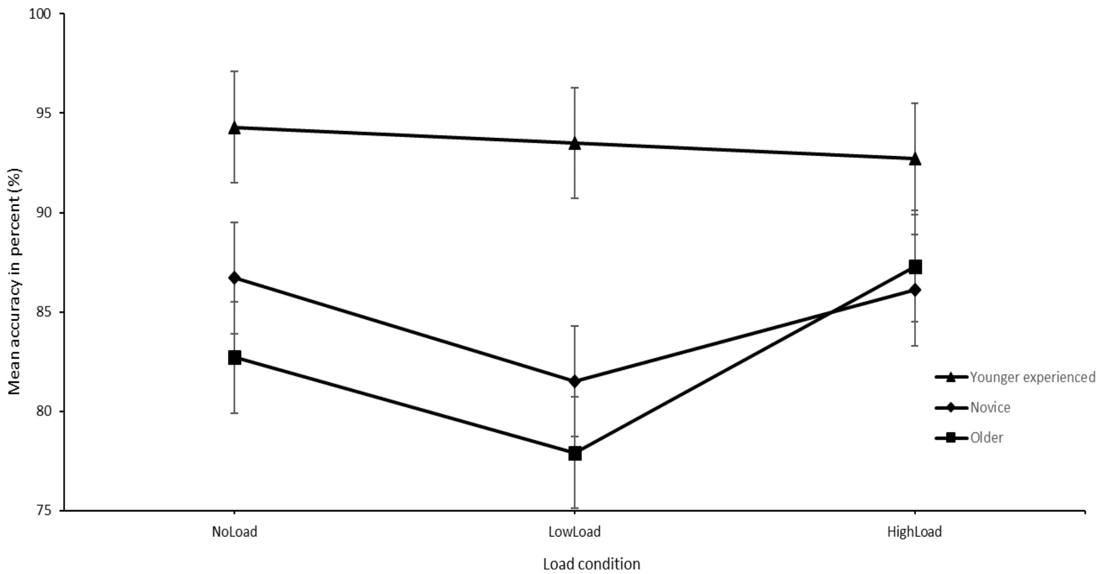

Figure 5.5. Mean hazard detection accuracy (±95% CI) as a function of load condition and driver group.

Inverse efficiency (IE).

As explained in Chapter 2, the inverse efficiency score combines reaction times and error rates in order to rule out potential trade-offs between speed and accuracy. This analysis revealed no main effect of group, *F(2, 87)* = 1.94, *p* =.15, η_p^2

=.04, or of load condition, *F(1.3, 115.9)* = .47, *p* =.62, η_p^2=.005. The interaction between load condition and group was also not significant, *F(2.7, 115.9)* = 1.31, *p* =.27, η_p^2=.03 (figure 5.6). Planned comparisons between young experienced and novice drivers at each load condition identified a significant difference in IE score between the driver groups during the low load condition (*p* =.02) but not during the no or high load conditions (*ps* >.06). Additional planned comparisons between young experienced and older drivers at each load condition found no significant difference in IA score at any of the load conditions (*ps* >.09).

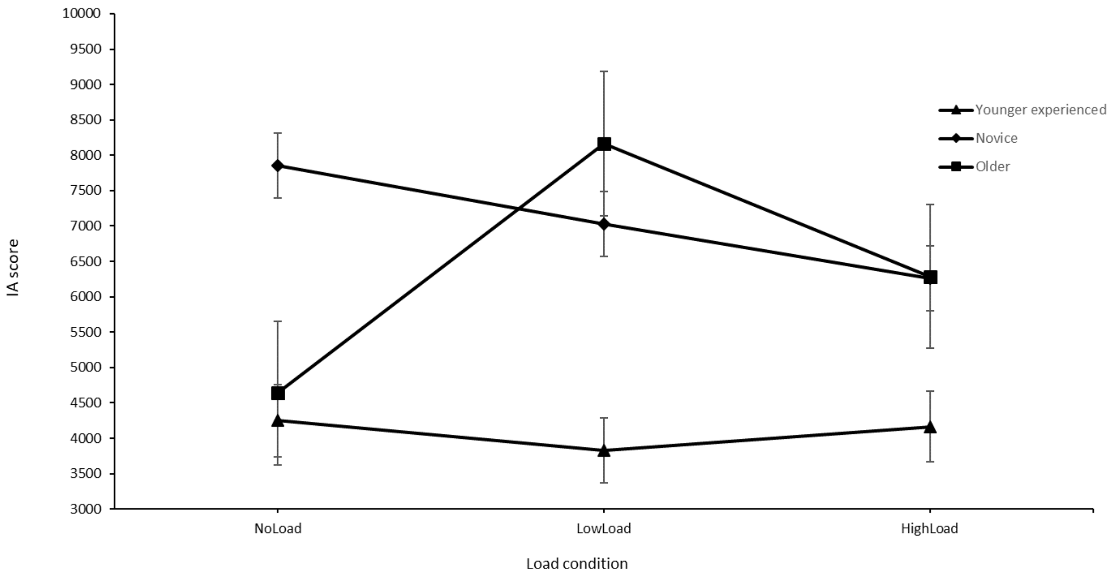

Figure 5.6. Mean IE score (±95% CI) as a function of load condition and driver group.

Missed hazard and incorrect identification of driving hazard.

One final consideration (which was not examined in earlier chapters) concerns the different types of error that participants can make in the hazard detection task; namely, either missing the hazard entirely, or incorrectly identifying the hazard. In the final analyses, I separated out these two types of error and calculated the proportion of error for each.

First, a mixed ANOVA was conducted on the missed hazard rates, using the within-subjects factor of load condition (no load, low load and high load), and experimental group as the between-subjects factor (young experienced, novice and older population, see Figure 5.7). In line with the analysis of overall accuracy, the effect of load condition was significant, $F(1.5, 131.1) = 6.22$, $p = .006$, $\eta_p^2 = .07$. Fewer missed hazard errors were made on average in the no load condition than in the low load condition (2.9% vs. 7.7%; $p = .006$). All other post hoc tests were non-significant ($p > .08$). The main effect of experimental group was not significant, $F(2, 87) = 2.19$, $p = .18$, $\eta_p^2 = .05$, however some differences between groups emerged in relation to the significant interaction between load condition and experimental group, $F(3.1, 131.1) = 2.79$, $p = .043$, $\eta_p^2 = .06$. Post hoc analysis showed that in the low load condition younger experienced drivers made fewer missed hazard errors than older drivers ($p = .02$), and in the high load condition younger experienced drivers made fewer missed hazard errors than novice drivers ($p = .04$). All other post hoc tests were non-significant $p > .06$. Planned comparisons between young experienced and novice drivers at each load condition showed that there was a significant difference in missed hazard errors between the groups during the high load condition ($p = .03$) but not during the no or low load conditions (*ps* >.09). Showing that younger experienced drivers made fewer missed hazard errors than novice drivers under higher load. This indicates that the skill of detecting hazards in situations of high cognitive load is better developed in the younger experienced drivers than in novice drivers. Additional planned comparisons between young experienced and older drivers at each load condition found a significant difference during low load ($p = .02$) but not under no or high load (*ps* >.10). Showing that younger experienced drivers made fewer missed hazard errors than older drivers under low load.

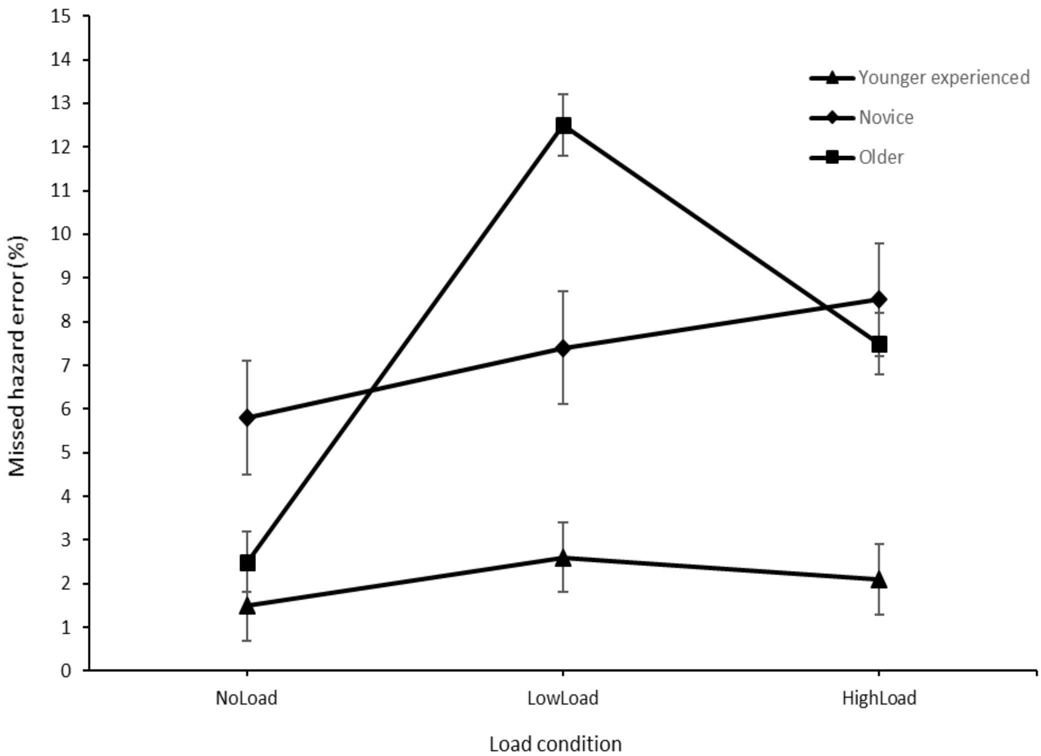

Figure 5.7. Missed hazard error rate (%) (±95% CI) as a function of load condition and driver group.

Next, a mixed ANOVA was conducted on the incorrect hazard identification rates, using the within-subjects factor of load condition (no load, low load and high load), and experimental group as the between-subjects factor (young experienced, novice and older population; see Figure 5.8). Interestingly, in contrast to the analysis of overall accuracy, the main effect of load condition was not significant, $F(2, 174)$ =1.69, p =.19, η_p^2 =.02. This suggests that the significant effect of load that was seen in the accuracy analysis was primarily driven by differences in missed hazard rates rather than incorrect identifications. This is perhaps unsurprising, because the missed hazard rates are likely to reflect trials in which the relevant hazard elements were not noticed and may therefore be more open to disruption by the concurrent workload of the non-hazard task than incorrect hazard identification rates (which are likely to reflect more of a misunderstanding of the situation and its severity). The main effect of experimental group was significant, $F(2, 87)$ = 5.4, p =.006, η_p^2=.11. Younger experienced drivers on average made fewer incorrect hazard identification errors than older drivers (4.2% vs 11%, p =.005). All other post hoc comparisons were non-significant (p >.17). The interaction between load condition and experimental group was also significant, *F(4, 174)* = 2.98, *p* =.02, η_p^2 =.06. In the no load condition, older drivers made more incorrect hazard identifications than both younger experienced drivers (*p* =.01) and novice drivers (*p* =.005). While this finding could reflect genuinely worse hazard perception in older drivers under low levels of workload, it could also be attributed to the fact that this group (unlike the two younger groups) will not have encountered hazard perception tasks as part of their initial driver training. In the low load condition younger experienced drivers made fewer incorrect identifications than both novice drivers (*p* =.03) and older drivers (*p* =.001). All other post hoc tests were non-significant (*p* >.07). Planned comparisons between young experienced and novice drivers at each load condition showed that there was a significant difference in incorrect hazard identification rates between the groups during the low load condition (*p* =.01) but not during the no or high load conditions (*ps* >.24). Additional planned comparisons between young experienced and older drivers at each load condition found significant differences under no load (*p* =.01), low load (*p* =.001) and high load (*p* =.04). This result shows that regardless of load condition older drivers are more likely to incorrectly identify a hazard than younger experienced drivers. This links back to the point above that older drivers will not have encountered hazard perception tasks as part of their initial driver training and therefore are disadvantaged with their lack of hazard perception training compared to younger experienced drivers.

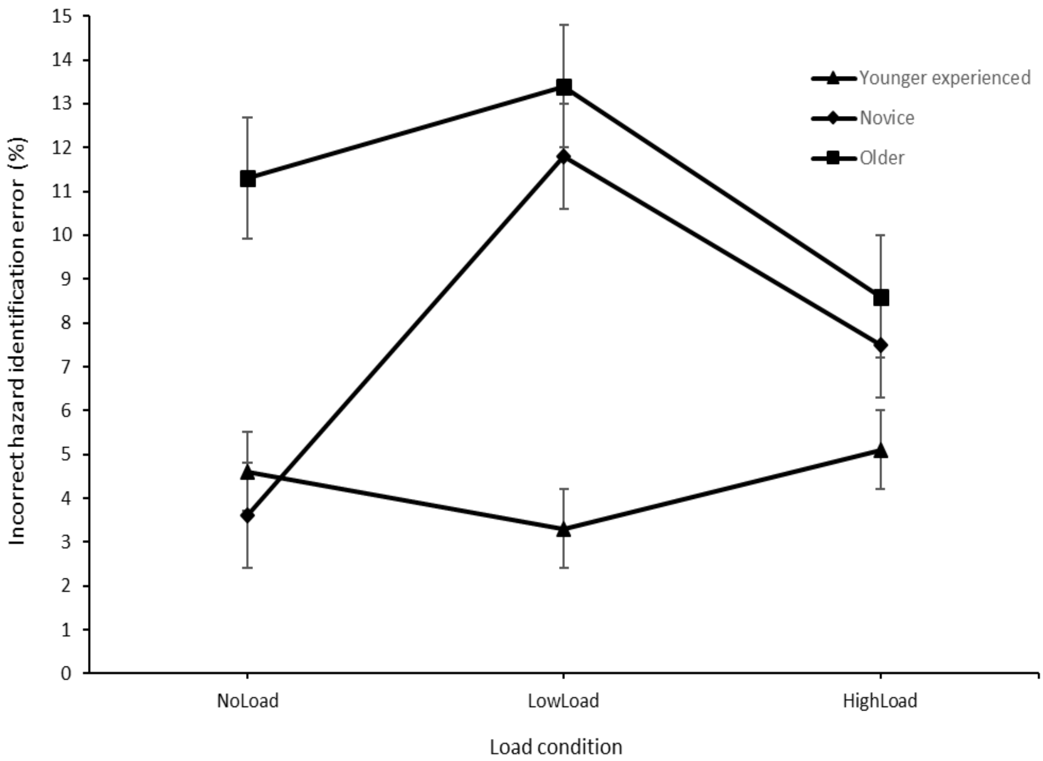

Figure 5.8. Incorrectly identified hazard error rate (%) (±95% CI) as a function of load condition and driver group.

Discussion

The detailed comparison of the three driving groups carried out in this chapter overall revealed a mixed pattern of results. RSME scores increased reliably between the load conditions and did not differ between the groups, indicating that the groups experienced a similar overall level of subjective workload. Yet, when looked at in more detail, older drivers in the no load condition rated the condition as requiring significantly more effort than the young experienced driving age sample reported. This might reflect an overall reduced capacity in the older drivers, such that they already experience the hazard detection task alone as relatively demanding. However, this could also reflect the fact that the older drivers are less experienced with hazard detection tasks, having not experienced them as part of their driver training (unlike the other participants). In addition, because of the subjective nature of this measure, it is also possible that older drivers simply use the ratings scale differently than the other participants.

However, the differences between older drivers and the other two groups persisted in the findings relating to accuracy and RT in the number probe task. Older drivers made more mistakes than both the young experienced and novice driving age samples and took significantly longer to respond to the probe. Although it at first seemed possible that older drivers may have been sacrificing their number task performance in order to maintain good performance on the hazard task, this seems unlikely because if anything their performance on the hazard task was on balance worse than that of the young experienced drivers, as I will now outline.

Overall, the factors of load condition and driving group had little effect on the RTs to detect the driving hazard. However, there was a small but significant difference between novice and older drivers under no load, such that novice drivers exhibited significantly longer RTs. This could relate to the relative inexperience of the novice drivers by comparison with the older drivers, who will have had exposure to a much wider variety of driving hazards. There was no difference in hazard detection accuracy between these groups, thus overall the hazard detection performance of the older drivers could be characterised as slightly better than that of the novice drivers under no load conditions. The fact that the RT difference only arose under the no load condition rules out the possibility (raised earlier) that the older drivers were sacrificing their performance on the number probe task in order to maintain good performance on the hazard detection task, because no such trade off would have been necessary under no load conditions where the number probe task was absent.

A different pattern of results was evident in the comparisons between the older and young experienced drivers. These groups did not differ in hazard detection RTs, but the older drivers did show worse overall accuracy than the young experienced drivers, as well as significantly lower accuracy under no load and low load conditions. Thus, the older drivers exhibited worse performance than the young experienced drivers, both on the hazard perception task and on the number probe task. This is in line with the wealth of literature showing worse hazard perception in older (vs. young experienced) drivers and, in some cases, worse dual task performance (Caird et al., 2005; Maquestiaux et al., 2013, 2010; Ross et al., 2012).

Interestingly, the novice drivers also performed worse overall than the young experienced drivers on the hazard detection task, because although their overall RTs were no different from those of the young experienced drivers, their overall accuracy was significantly worse. In addition, under low load conditions, the novice drivers' RTs and accuracy were both significant worse than those of the younger experienced drivers. This is also in line with the substantial research evidence demonstrating worse hazard perception in novice (vs. more experienced) drivers (Borowsky & Oron-Gilad, 2013; Horswill et al., 2010; McDonald et al., 2014; Scialfa et al., 2012).

It is important to note that, although the above patterns of results are identifiable when comparing between the RT and accuracy analyses, almost no group differences emerged when these measures were combined into a single inverse efficiency score (with the exception of the novice drivers showing worse performance than the younger experienced drivers under low load). However, the inverse efficiency measure is open to criticism. For example, the use of IE scores can increase the variability of the data particularly if the study has a small number of observations (<20) per condition (as was the case in this study, in which 20 observations were collected per load condition) and in cases where there is little correlation between RT and accuracy (Bruyer & Brysbaert, 2011). However, these authors also argue that the use of IE scores is acceptable as a complementary analysis to RT and accuracy analyses, and thus the IE measure was only used in this context in the present work.

Chapter 6 - The Effect of Automation on Takeover Performance and Situation Awareness

Introduction

The experiments described so far in this thesis used an abstract number monitoring task at different levels of load as a proxy for the different levels of driving demand that will result from differing levels of in-car automation. Although this approach has the advantage of allowing a high level of experimental control, it also suffers from a significant lack of ecological validity, because the real task of manual driving is of course far more complex than even the high load version of the number monitoring task used in the work reported so far. For this reason, the current chapter reports a study in which I used a driving simulator to achieve much more realistic levels of task demand.

In addition, whereas Chapters 2-5 of this thesis relate to fairly low levels of automation (in which drivers are required to monitor the driving environment even when the autonomous system is engaged), the current study focuses on a higher level of automation, in which drivers can fully disengage from the task of driving during periods of automated function. This is because the industry is already moving very quickly towards these higher levels of automation, making the lower levels increasingly less relevant for current research. For example, one of the current market leaders in the development of autonomous driving features, Tesla (USA), has already created an 'advanced autopilot' mode that can match speed to traffic conditions, keep within a lane, automatically change lanes without requiring driver input, transition from one roadway to another, exit the roadway when the driver's destination is near, self-park when near a parking spot and be summoned to and from the driver's garage. The ability to have all these tasks completed by the autopilot system greatly reduces the demands of driving for the driver. However, the system requires the driver to maintain the ability to take back control of the car if the system requires the driver to do so. This requirement of drivers to regain control safely and in a timely manner (i.e. return to the driving 'loop' from being 'out of the loop'), often having been engaged in a non-driving-related secondary task while 'out of the loop' (e.g. using phones, laptops, reading or writing emails), is a critical issue for research as more level 2/3 autonomous systems enter the mainstream vehicle market. This is the focus of the current study.

Key Aims

The key aims of this study are to examine how drivers regain control of a vehicle after a period of autonomous driving, where their attention is focused during this handover process, and how quickly the attentional and behavioural patterns of driving return to a level like that of manual driving. In addition, many vehicle designs are now allowing for a complete shift of driver body position during periods of automation, so that drivers can engage with non-driving tasks in an entirely different section of the vehicle (e.g. pivoting in the driving seat to use a laptop station that is mounted in between the driver and passenger seats). In comparison with designs that keep the driver facing forwards in a static driver seat, these more flexible designs will reduce the ongoing level of situation awareness that the driver is able to maintain, because in some cases they will be able to readjust their body position such that very little relevant visual information is available (e.g. by spinning the seat through 180 degrees to face backwards). Thus, an additional research question in the current study concerns the impact on the handover process of the level of visual information from the driving scene that has been available throughout the automated period. Rather than manipulating this through changing the driver's body position (which would have been practically challenging, particularly in relation to the collection of eye movement data), I instead introduced heavy fog in some conditions, reducing the amount of visual scene information to close to zero.

Takeover Time Following Automated Driving

There is a large amount of existing research examining the time needed to take over manual control after a period of automation. Eriksson and Stanton (2017) reviewed existing research into driver control transitions to determine the time it takes drivers to resume control from an AV in noncritical scenarios. Their initial meta-analysis of 25 relevant studies highlighted that takeover-request lead times (TORlt) and takeover RT (TOrt) varied widely from study to study with times varying from 0 to 30 secs for TORlt and 1.14 to 15 secs for TOrt. Their review highlighted that the mean TORlt was 6.37 sec with a mean TOrt of 2.96 secs. The most commonly used TORlt was 3 secs with a mean TOrt of 1.14 secs. It is important to note the significant variability identified between the results of studies examining the same phenomenon (i.e. takeover request RT). This is likely to be driven by differences in TORlts and other methodological differences.

Indeed, one recent simulator study directly compared the effects of three different TORlts. Mok, Johns, Miller, and Ju (2017) measured the driving performance of 30 participants at three abrupt transition times (2, 3 and 8 secs before the vehicle entered a sharp corner with reduced lanes). The secondary task performed during automated driving was playing a simple computer game on a tablet computer. The driving task was simply to disengage from the task and regain control of the car when the transition from automation to manual driving occurred. Interestingly, most drivers were observed to drop the tablet on the passenger seat or on their lap (28/30) in response to an urgent handover request. Among the more formal performance measures were the standard deviation of steering wheel position. This measures quality of driving performance, with a smaller deviation indicating finer vehicular control. These results showed that there was a significant difference between the 2 second (Mean = 1.29), the 5 second (Mean = 0.48) and the 8 second (Mean = 0.19) transition times, such that the longer transition times had less variability and thus showed finer control than the shorter transition times. The standard deviation of lane position (SDLP) results showed the same pattern, with a significant difference between the 2 second (Mean = 1.88), the 5 second (Mean = 0.6) and the 8 second (Mean = 0.28) conditions. In summary, the shortest transition time produced a coarser and less smooth set of driving behaviour responses compared with the longer transition times. In the light of this previous research, the current study used an intermediate TORlt of 3 seconds, which is in line with the majority of the published experiments.

Previous research on this topic has also identified significant variability between participants in the time taken to regain control. For example, Eriksson and Stanton (2017) also carried out a simulator-based study in which 26 participants completed two scenarios with an autonomous driving system activated. Drivers were either asked to read a newspaper, or to monitor the system, and to relinquish, or resume, control from the autonomous system when prompted. The results showed that drivers took between 1.9 and 25.7 seconds (median = 4.6 secs) to resume control from automated driving under normal conditions but when drivers were engaged with a secondary task the time taken increased to between 3.2 and 20.9 seconds (median = 6.1 secs). These findings demonstrate that there are likely to be individual differences between drivers' RTs and reengagement after automation, which will be an important consideration in evaluating the findings of the current study. In addition, because participants in the current study will be engaged in a secondary task during the automated period, I expect them to be slower at regaining control than participants in studies where no secondary task is completed.

Situation Awareness Following Automated Driving

One of the central questions of the present study concerns people's visual attention (as measured by their eye movement behaviour) following the handover of an automated driving system, as well as their subsequent responding to developing hazards in the driving environment. The clear prediction is that the disengagement of the driver during the period of autonomous driving will have significant effects on situation awareness when manual driving is subsequently reengaged (by comparison with conditions of continuous manual driving without any period of autonomous control). Although to my knowledge no studies have looked at exactly this question, there is evidence that differing levels of engagement during autonomous driving lead to differences in situations awareness when the system subsequently hands over control. For example, Louw, Madigan, Carsten and Merat (2016) ran a simulator study examining drivers' patterns of fixation during partially-automated (SAE Level 2) driving on approach to critical and non-critical events. 75 participants completed five different conditions which used different levels of screen visibility and secondary tasks to induce varying levels of engagement with the driving task. The 5 conditions were: no fog, light fog, heavy fog, heavy fog plus a visual quiz task, and no fog plus a simple cognitive task. Drivers' first fixations back to the screen after automation disengagement were influenced by the condition. In the no fog, light fog and heavy fog conditions, between 55-65% of the first fixations fell in the centre of the screen (a 6^{o} circle in the centre of the screen). However, in the two conditions with a secondary task, a lower proportion of first fixations fell in the centre of the screen (35-40%), indicating a more diffuse pattern of first fixations in these conditions. The percentage of fixations falling in the road centre (PRC) was calculated for three one-second time windows, immediately after the drivers were prompted to take back control of the vehicle. PRC is defined as the proportion of fixation data points which fall within the road centre area, a 6^{o} circular region located around the driver's median fixation point. PRC has previously been demonstrated to be a sensitive indicator of visual distraction with lower values indicating less attention is dedicated to the visual demands of driving (Victor, Harbluk, & Engström, 2005). Regardless of condition, the PRC in the first second was significantly lower than in the second (65% vs. 75%) and the difference between the 1^{st} and 3^{rd} second (65% vs. 78%) was also significant. Importantly, in the two conditions with the highest visual occlusion of the screen (heavy fog and heavy fog plus task) the PRC was significantly lower in the 1^{st} second compared with the other conditions but then recovered to a similar level as to the other conditions. This potentially indicates that drivers were scanning the driving environment more diffusely after not having seen the road beforehand, but at the expense of focusing on the centre of the screen. These findings might at first suggest that participants' visual attention in the current study will be more diffuse following the conditions in which the driving environment is occluded by fog (vs. conditions in which no fog occurs). However, an important difference between the current study and that of Louw et al. is that participants in my study will be looking away from the driving environment during the automated period, engaging with a tablet or phone. Thus, any performance reductions in the fog (vs. no fog) condition in the current study would suggest that drivers were making some use of the visual information available in the no fog condition, despite not explicitly directing their attention towards the screen.

Recovery from Automation Failure

A key question in the study of automation concerns the driver's ability to resume control from an automated system, particularly after a failure in automation. There is evidence to suggest that the unexpected disengagement of an automated system can have a detrimental impact on vehicle control and visual attention. For example, in a recent driving simulator study, Merat, Jamson, Lai, Daly and Carsten (2014), examined drivers' ability to resume control from a highly automated vehicle in two automation conditions. The first condition examined when automation was switched off and manual control was required at regular intervals. In the second condition, the transition to manual occurred when the participant had been looking away from the road for 10s or more. Performance, measured by steering behaviour and fixations towards the centre of the road, was significantly better in the first condition compared to the second condition. Specifically, visual attention towards the driving scene was erratic for up to 40s after the less predictable transfer of control, compared to when disengagement was predictable and at a fixed pace. Additionally, in the second condition following the unpredictable transfer, driving measures and eye fixations showed a 10-15 second lag in time between disengagement of the automation and resumption of control by the driver, with the largest discrepancy in the first 15 seconds following transfer, compared to more stabilised vehicle control after 35-40 seconds. These results demonstrate that there are significant potential costs to driving performance of the unexpected disengagement of an automated system.

Research has also suggested that the extent of the performance costs associated with automation failures relates to the level of automation experienced before the failure. For example, in a study using a high-fidelity driving simulator, Strand, Nilsson, Karlsson and Nilsson (2014), examined semi-automated (ACC) and highly automated driving (HAD) conditions in which there were 3 levels of system failure (complete, severe, and moderate failures). They found that drivers who were engaged in HAD and experienced any system failure were involved in more near collision events than those driving using a semi-automated vehicle (43% vs. 22%). The effect of system failure on the number of near collisions between the HAD and ACC conditions was only significant in the complete system failure condition. Additionally, Radlmayr, Gold, Lorenz, Farid and Bengler (2014) using a high-fidelity driving simulator compared automation failure response times in a HAD condition to a manual driving response times. The manual driving condition was a baseline with no automation failure but they experienced the same track and driving hazards as the HAD condition. In the HAD condition response times were longer than the baseline manual condition (2370 vs. 1850 milliseconds), although the number of near collisions was comparable across both condition (seven vs. five, N=24 participants in each condition).

However, alternative evidence suggests that the level of automation that is engaged before the transfer of control might not have a large effect on driving performance following unexpected transfers of control. For example, Young and Stanton (2007) demonstrated that level of automation had no influence on the braking response time (BRT) following failure of automation. Using a driving simulator the level of automation was varied across two conditions: one with ACC and a second with ACC and active steering (AS). AS is a lane-keeping device, which keeps the vehicle in the centre of its lane until the system is disengaged or manual steering input is received from the driver. BRT was then measured in recovering control of the vehicle following failure of automation and no difference was observed between the two automation conditions. When the data of the two automation conditions were compared with previous experiments on manual driving BRT, response times were substantially longer (1000-1500 ms) when using automated (vs. manual) systems. However, the lack of a manual comparison group in the original experiment makes it difficult to draw firm conclusions on the basis of these findings, because they rely on comparisons between two experiments with different methodologies. Nevertheless, the results do suggest that a reduction in task specific workload – namely, the need to control the vehicle’s heading and speed - rather than improving task performance conversely might interfere with recovering control of the vehicle following failure of automation. A plausible explanation for this decrease in performance is that reducing mental workload, creating cognitive under-load, is not necessarily a good thing, particularly in cases where the task is already manageable. Indeed, this mental under-load has been claimed in some circumstances to be as detrimental to performance as very high workload (Desmond & Hoyes, 1996; Young & Stanton, 2002). Recall, in addition, that similar effects were seen in Chapter 2, such that participants performed worse on the hazard detection task under lower (vs. higher) workload conditions.

Study Rationale and Hypothesis

In this study, I set out to examine the effect of reduced visual information during a period of automated driving on drivers' subsequent situation awareness and resumption of vehicular control. To examine this, I used a screen manipulation that reduced the amount of visual information available during automated periods and measured drivers' eye movements and driving behaviour in the time following a handover request. My study considered the following questions:

i. What is the time course of driving behaviour after the resumption of control following a period of automated driving?
ii. Does the amount of visual information that is available during automated driving influence drivers' eye movements and driving behaviour when they resume control?

Method

Participants

Forty participants were recruited for the study (mean age = 49.5, SD = 11.1, 23 males). All participants held a full UK driving licence (mean years licence held = 31, SD = 11.4) and drove more than 3 times a week (mean annual mileage = 13,750, SD = 9075). In addition, all participants self-reported normal or corrected-to-normal visual acuity (normal = 12 participants, corrected-to-normal = 28 participants) as well as normal colour vision. Two participants had experienced driving automation of level 2 or above (Tesla autopilot and GPS controlled farm machinery). All participants had previously taken part in studies at TRL and experienced the driving simulator. Participants were paid £30 for their participation.

Apparatus

Visual stimuli were presented on three 27-in monitors (Asus MX279H) each with a resolution of 1920x1080 pixels, a refresh rate of 60Hz and a viewing distance of approximately 90cm to the central monitor. Eye movements from the central monitor were recorded using a Tobii X-1 Light running at 20hz. A nine-point calibration procedure was used and accepted only if the average error was less than 0.5° of visual angle and the maximum error was less than 1.0° of visual angle. Head position was unrestricted. Auditory stimuli were presented through the stereo speakers present in the monitors at a comfortable listening volume (approximately 60dB). Participants controlled the car using a steering wheel, clutch, brake, accelerator and gear stick attached to the simulator computer (Logitech G27 Racing Wheel). (Figure 6.1).

Figure 6.1. Example of the simulator set up (note eye tracker not included in this photo)

Stimuli

Driving environment.

Participants were presented with a simulated motorway driving environment for the duration of the study (see Figure 6.2). The simulator itself was implemented in SCANeR II studio (www.oktal.fr) with a data sampling rate of 20Hz, and a custom eye tracking module interface designed to integrate with the SCANeR II software and the Tobii. The driving environment presented in this study was a three-lane motorway with gentle S-curves, with intermittent overhead gantries and occasional bridges. The motorway had a maximum advised speed limit of 70 mph (112 kph). The landscape was flat and wide open with trees in the surrounding area. The two carriageways were separated by a double crash barrier in the middle. There was no traffic presented on the opposing carriageway. The ego vehicle (vehicle driven by participants) showed a 1st person view of the driving environment and had the speedometer, gear and rev counter presented on the bottom of the screen in the centre. The wing mirrors were presented on the left and right monitors, in the top left corner of the screen for the left monitor and the top right for the right monitor; there was no rear-view mirror. The motorway environment the driver drove through consisted of 50 vehicles which swarmed around the ego car within the following parameters: 1) the average headway of the swarm vehicles was 1.5-2 secs with an average speed of 65-70 mph (104-112 kph); 2) swarm vehicles were distributed between all three motorway lanes with the majority in the centre and left lane; 3) when swarm vehicles were over 700 metres in front of the ego vehicle they either regenerated 600 metres in front of the ego vehicle or 100 metres behind. This allowed the driving environment to remain busy and relatively crowded regardless of each driver’s individual driving style and speed, but not create an environment where the driver perceived ghost vehicles (disappearing and appearing vehicles). Finally, swarm vehicles could change lane if changes to the driving environment required them to.

Figure 6.2. An example of simulated motorway driving environment used in the study (note the speedometer and rev counter are not shown in this image).

Driving hazards.

I created in the simulator software two common types of driving hazard that can occur in the motorway environment: a crash between two vehicles blocking a lane, and cones reducing the carriageway down from three lanes to two. Both hazards occurred 150 metres after the ego car passed a programmed data trigger, making them just visible in the distance to the driver. The crash obscured the left lane and involved a collision between a small lorry and a car, forcing all swarm vehicles in the left lane to move into the centre lane. The cones hazard again reduced the carriageway from three to two lanes, and forced the swarm drivers to move from the left to the centre lane. The cones were spaced at 7 metre intervals according to department for transport guidelines (Department for Transport/Highways Agency, 2009) for 700 metres.

Autonomous system.

The autonomous driving system was a custom module created by TRL. The autonomous system was programmed to maintain an average speed that matched the roadway (e.g. 65-70mph (104-112 kph) for motorway) at a minimum headway of 1.5-2 secs, and the speed at the point of takeover was 65-70mph (104-112 kph). It maintained lane position and controlled lateral position in the lane. The handover and takeover requests both consisted of a 3 second sequence of auditory tones which was matched as closely as possible to the tones used in BMWs and Teslas. The sequence consisted of three 250ms tones (700Hz, 44.1khz sample rate) equally presented over 3 seconds with the final tone the point at which either the autonomous system took over driving or handed back control to the driver.

Screen manipulation.

I used a screen manipulation to assess the impact of reduced visual information during automation on the resumption of control after a period of automation. The two-screen conditions were: no screen manipulation; and fog screen manipulation. Both screen manipulations were only active when the autonomous system was driving the ego vehicle. The fog screen manipulation was a heavy fog that obscured 90% of the visual information on the screen, bar faint road markings within 1-2 metre of the ego vehicle. The fog screen manipulation was removed within 10ms of the first takeover tone occurring.

Design and Procedure

Participants were asked to drive the simulated vehicle safely through the driving environment whilst controlling the lateral and longitudinal direction with the steering wheel, clutch, brake, accelerator and gear stick. They were asked to maintain their speed at approximately 70mph and to stay in the centre lane at all times unless there was a serious risk of them crashing into swarm vehicles (at which point they could move into the right lane). They were informed that in some trials they would handover control of the vehicle to an autonomous driving system (as per the auditory tones) and that they would be required to resume control of the vehicle when requested by the system (also as per the auditory tones). Participants completed one two-minute practice trial to familiarise themselves with the task, the handover/take back request and driving controls. There then followed eight experimental trials (manual driving leading to crash hazard, manual driving leading to cones hazard, autonomous drive with no hazards and no fog, autonomous drive with no fog leading to a crash hazard, autonomous drive with no fog leading to a cones hazard, autonomous drive with fog but without any hazard, autonomous drive with fog leading to a crash hazard, and autonomous drive with fog leading to a cones hazard). I counterbalanced the trial order to ensure that each trial did not occur more frequently in one of the eight possible trial positions across participants.

The manual and autonomous drive conditions varied slightly in their procedures. In the manual condition, the drivers drove for approximately 2 minutes before they reached the data trigger at which point one of the two hazards occurred. After a further two minutes the trial ended. In the autonomous drive, the participants drove for approximately 2 minutes before the handover request occurred. The autonomous system then drove for 7-8 minutes, during which time participants were encouraged to engage with their mobile phone or tablet, doing whatever they would normally do on their device. After 7-8 minutes, the takeover request occurred. Approximately 3 seconds after the end of the takeover request the participants passed the hazard data trigger and one of the three hazard conditions occurred (no hazard, crash hazard or cones hazard). Approximately 3 minutes after passing the hazard data trigger the trial ended. The entire experiment took 90 minutes to complete, including the practice trials.

Results and Preliminary Discussion

In the results below, I begin by describing our analytic approach for this study, and the manner in which the data were prepared for analyses. I then describe the results of those analyses in relation to a series of dependent measures – six behavioural measures of driving performance as well as three eye tracking measures. The first of the measures of driving behaviour concerns the time taken for participants' hands to return to the wheel. This provides a measure of motor readiness to resume control of the vehicle after the period of automation. Next, I examined braking RT to the driving hazard to assess drivers' situation awareness in relation to developing hazards. I also examined the participants' speed, headway to the vehicle in front and time to contact with the vehicle in front, all measured from the start of the developing hazard. The final behavioural measure was the time course of lane position of the drivers over 60 seconds, to ascertain their resumption of control after automation. In the eye tracking analysis, I examined the first fixation time back to the screen at the start of the takeover request, to ascertain how quickly drivers were able to set aside their secondary task (engaging with phone or tablet) and redirect their gaze back to the driving scene. I also examined the time course of the PRC over 60 seconds following the handover, to examine the extent to which drivers directed their visual attention to the central area of the screen during that period. Finally, I examined first fixation durations to determine whether there were any basic processing differences between the different conditions. Because of the complexity of the analyses and the number of dependent measures, I provide preliminary discussion of the results from each measure as it is presented.

Analytic Approach

Eye tracking studies in which dynamic, moving displays are used produce a qualitatively different form of eye movement data than studies which use static displays. Unlike static displays, where participants tend to make fixations of relatively short duration (e.g. 200-250ms), in dynamic displays like those used here, participants often make longer, 'smooth pursuit' fixations, during which the fixation position moves slowly enough to enable the acquisition of visual information without making a saccade (which would prevent the acquisition of visual information). It has been argued (e.g. Baayen, Davidson, & Bates, 2008; Hua Huang, Ming Luh, Sheu, Tzeng, & Chen, 2011) that Mixed Linear Models (MLMs) constitute a more appropriate analysis method for data of this type than the more standard approaches (e.g., *F-tests*, standard *t*-tests) for several reasons. First, the MLM approach is able to capture the full variability of the dataset, because MLMs examine data pertaining to each individual data point, rather than mean-averaging the data as is the case with standard statistical tests. This is important since I am analysing data derived from a dynamic, changing environment so the data were naturally different than would be observed in a static task. Second, this approach has the advantage that participants can be added as a random factor to the models, and the resulting models can shift their fits based on each individual participant. As a result, the variable strategies or methods adopted by each participant to complete the driving task can be captured, to a certain extent, by allowing the model to modify its fit based on each participant. Mixed-effects modelling treats participants and items (if specified) in an experiment as random samples from larger populations (Raaijmakers, Schrijnemakers, & Gremmen, 1999), whereas ANOVA analysis fails to look at the systematic variability due to individual items, and overlooks the systematic variability due to individual participants (Locker, Hoffman, & Bovaird, 2007). For these reasons, MLMs are becoming a more and more common analysis method in eye tracking studies (e.g. Godwin et al., 2015; Godwin, Benson, & Drieghe, 2013; Godwin et al., 2013; Laubrock, Engbert, Rolfs, & Kliegl, 2007).

Data Preparation

The data were prepared in the following manner. For analyses that related to participants' responses to takeover requests (which only involved autonomous conditions), I used 60 seconds of the data indexed to 0 seconds from the start of takeover request (the beginning of the first auditory takeover request tone). This allowed me to analyse in detail any specific differences and time course of any differences in takeover performance between the autonomous conditions. For analyses that related to a hazard response (including both autonomous and manual conditions) I used 60 seconds of the data indexed to 0 seconds from the start of hazard data trigger. This allowed me to compare any driving performance and situation awareness differences in all conditions in an equal and consistent manner. Further data preparation will be explained for each dependent variable below. A data recording error meant that all manual crash hazards conditions were not analysed in the subsequent sets of analysis.

Behavioural Data Analysis

Hands returning to the wheel.

Video of each experimental run was recorded on a Go-pro camera which was placed to the right of the participant and covered a view that encompassed the wheel and the participants' hands (but not their face). I used these videos to manually calculate the time it took participants to return both of their hands to wheel in all autonomous conditions, measured from the start of the takeover request. A repeated measures ANOVA was conducted using the DV of hands to the wheel RT and IV of screen manipulation (no-fog, with-fog). There was a significant effect of screen manipulation on hands to the wheel RT, $F(1, 39) = 6.94$, $p = .012$, $\eta_p^2 = .15$, power =.73, such that responding was slower in the no-fog condition (4150 ms) than in the fog condition (3850 ms), see Figure 6.3.

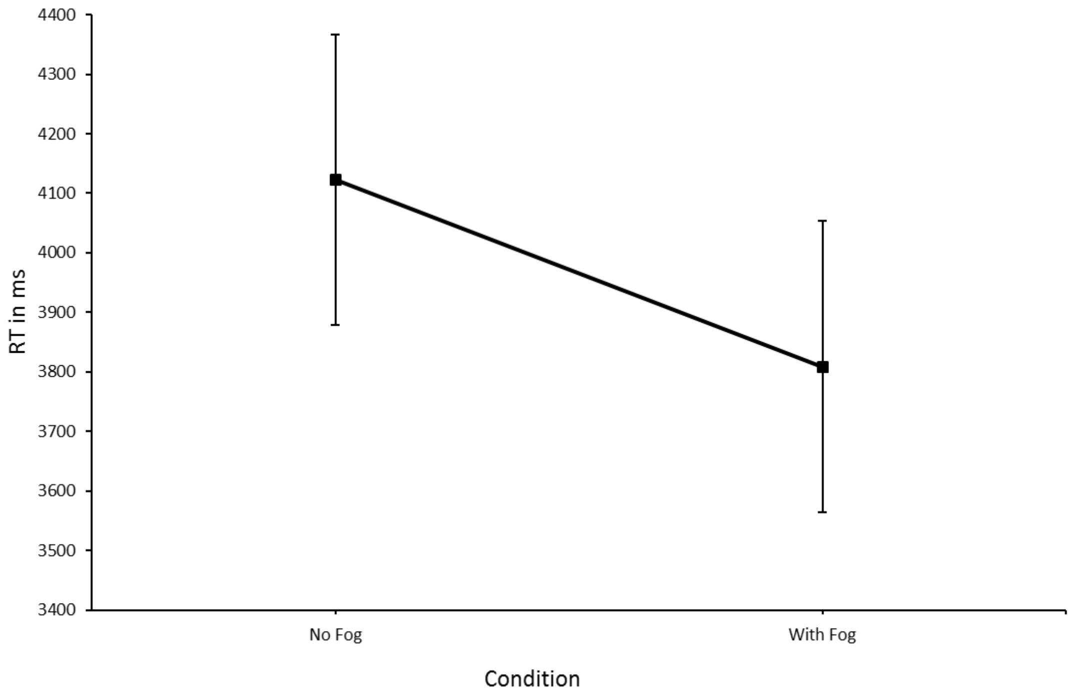

Figure 6.3. Time taken to return hands to the wheel in milliseconds (error bars represent standard deviation)

These results have two interesting implications. First, when drivers have been engaging with their phones or tablets in a lifelike manner (as they were in the current study), it takes them around 4 seconds on average to return their hands to the wheel following a takeover request from an autonomous system. Second, people were slower to return their hands to the wheel in the ‘no fog’ condition, where visual information about the driving scene remained available during the automated period (vs. the ‘fog’ condition where no such information was available). Although at first this may seem to be a surprising result, it is possible that the abrupt removal of the fog (which occurred at the start of the takeover request) led to an alerting effect, speeding reaction times in that condition. If this is the case, a similar effect should be seen in the eye tracking analysis (such that first fixations back to the screen should also be faster in the fog vs. the no fog condition).

Braking RT.

Comparison of autonomous conditions (fog vs. no fog).

The braking RT was calculated, measured from the onset of the developing hazard, as a function of the screen manipulation (fog vs. no fog). A MLM was conducted on the braking RT using the lme4 package in R (Version 3.4.0; Bates, Maechler, Bolker, & Walker, 2017). All reported *p*-values were obtained using the lmerTest package in R (Kuznetsova, Brockhoff, Haubo, & Maintainer, 2013). Linear mixed model fit by REML (restricted maximum likelihood) t-tests use Satterthwaite approximations to degrees of freedom from the lmerTest package, which produces acceptably low Type 1 error rates even for smaller samples. I conducted contrasts to explain main effects within the models using the lsmeans R package (Lenth, 2016), correcting for multiple comparisons where required. For the model participants was used as the random factor and screen manipulation (no-fog, with-fog) was the fixed factor. Marginal R^2 represents the variance explained by the fixed factors and conditional R^2 represents variance explained by both the fixed and random factors (i.e. the entire model).

There was no significant difference in braking RT between the two screen manipulations: No fog (mean= 3.81sec) vs. fog (3.61sec), $t(103.67)$ =0.47, p =.64, Marginal R^2 = .001, Conditional R^2 = .08. Thus, by the time the hazard started, there were no effects of the amount of visual information that had been present during the autonomous period. Given that the hazard started at approximately 3-4 seconds after the handover, this finding suggests that, even if there were any initial differences in situation awareness between the fog and no fog conditions (a possibility which is examined further in the following analyses), these were not detectable in hazard braking responses after 3-4 seconds.

Comparison of manual condition with autonomous condition (no fog).

A second model was constructed using participants as the random factor and including all of the different trial types (manual cones hazard, autonomous drive crash hazard no fog, autonomous drive cones hazard no fog, autonomous drive crash hazard with fog, autonomous drive cones hazard with fog) as the fixed factor. There was a significant overall effect of trial $t(142.9) = 2.29$, $p = .02$, Marginal $R^2 = .07$, Conditional $R^2 = .10$.

The most important post-hoc comparison in relation to the study's main hypotheses is between the 'manual cones hazard' condition and the 'autonomous drive cones hazard no fog' condition. This allows an examination of hazard responding following the handover of an autonomous system, compared with hazard responding following a period of manual driving. (Recall that the data from the 'manual crash hazard' condition were not usable due to a technical error, meaning that the cones hazard is the only condition suitable for this analysis). This comparison was not significant ($p > .15$), suggesting that braking in response to a developing hazard was no slower following an autonomous handover than following a period of manual driving.

Results from remaining trial comparisons.

For completeness, this section reports the results from the remaining post-hoc contrasts. These showed a significant difference between hazard braking RTs in the 'autonomous drive cones hazard no fog' condition (4.6 sec) and the 'autonomous drive cones hazard with fog' condition (4.5 sec) ($p = .02$). Interestingly, this replicates the finding from the 'hands to the wheel' analysis, with faster responding in the 'fog' versus the 'no fog' condition. However, it is important not to overstate this finding, because recall that the earlier analysis found no significant difference in braking RT as a function of screen manipulation overall. There was also a significant difference between the 'autonomous drive cones hazard with fog' condition (4.5 sec) and the 'autonomous drive crash hazard with fog' condition (2.5 sec; $p = .01$) indicating that within the 'autonomous drive with fog' conditions, a faster reaction time was obtained to the crash hazard than to the cones hazard. All other contrasts were non-significant (p >.15).

Speed at developing hazard (kph).

Comparison of autonomous conditions (fog vs. no fog).

The speed at which participants were driving at the start of developing hazard was analysed in the same way as the braking RTs. For the first model, participants was used as the random factor and screen manipulation (no-fog, with-fog) as the fixed factor. In line with the analysis of the braking RTs, there was no significant difference in speed at the start of the developing hazard between the no fog condition (mean= 90.7 kph) and the fog condition (92.1 kph), $t(157.9)$ =0.77, p =.44, Marginal R^2 = .003, Conditional R^2 = .02.

Comparison of manual condition with autonomous condition (no fog).

As in the braking RT analysis, a second model was constructed using participants as the random factor and including all trial types (manual cones hazard, autonomous drive crash hazard no fog, autonomous drive cones hazard no fog, autonomous drive crash hazard with fog, autonomous drive cones hazard with fog) within the fixed factor. There was an overall significant effect of trial $t(269.6)$ =2.71, p =.007, Marginal R^2 = .11, Conditional R^2 = .14.

Once again, the most theoretically interesting post-hoc comparison is between the ‘manual cones hazard’ condition and the ‘autonomous drive cones hazard no fog’ condition. In contrast to the braking RT analysis, the speed at the onset of the hazard did differ between these two conditions. The post-hoc contrast showed a significant difference in approach to the cones hazard between the manual condition (101.3 kph) and the autonomous no fog condition (88.2 kph; p <.001). Interestingly, this indicates that participants were driving more slowly at the time of the hazard onset following handover of the autonomous system than following a period of manual driving, perhaps suggesting that drivers are more cautious in the period following the handover and therefore adopt lower speeds.

Results from remaining trial comparisons.

For completeness, I note that the post hoc contrasts also revealed differences between the 'manual cones hazard' condition (101.3 kph) and the 'autonomous drive cones hazard with fog' condition (88.1 kph; p <.001). However, this comparison is not particularly meaningful because it confounds the effects of manual (vs. autonomous) driving with the effects of fog (vs. no fog). There was also a significant difference between the 'manual cones hazard' condition (101.3 kph) and the 'autonomous drive crash hazard with fog' condition (91.7 kph) but again this comparison is not meaningful because it confounds driving condition (manual vs. autonomous), fog presence and hazard type. All other contrasts were non-significant (p >.09).

Headway to lead vehicle.

Comparison of autonomous conditions (fog vs. no fog).

The headway in metres to the lead car at the start of the hazard was analysed in the same way as the previous two measures. For the model, participants was used as the random factor and screen manipulation (no-fog, with-fog) was the fixed factor. In line with the braking RT and speed results, no significant difference was found between the no-fog condition (111.9 metres) and the fog condition (113.1 metres), $t(179)$ =.16, p =.88, Marginal R^2 = .0001, Conditional R^2 = .0001.

Comparison of manual condition with autonomous condition (no fog).

As in the preceding analyses, a second model was constructed using participants as the random factor and trial type (manual cones hazard, autonomous drive crash hazard no fog, autonomous drive cones hazard no fog, autonomous drive crash hazard with fog, autonomous drive cones hazard with fog) as the fixed factor. There was a significant overall effect of trial $t(207.1)$ =2.48, p =.01, Marginal R^2 = .07, Conditional R^2 = .08. Post-hoc contrasts showed no significant difference between the 'manual cones hazard' condition and the 'autonomous drive cones hazard no fog' condition (p >.05). Thus, although the previous analysis indicated that drivers were travelling more slowly at the hazard onset following a period of automated (vs. manual) driving, perhaps indicating a more cautious approach, this was not mirrored in the amount of headway that they left in relation to the car in front.

Results from remaining trial comparisons.

For completeness, I note that the contrast between the 'autonomous drive crash hazard no fog' condition (133.2 metres) and the 'autonomous drive cones hazard with fog' condition (93.2 metres) was significant (p =.03). However, this comparison is not meaningful because it confounds hazard type and fog presence. No other contrasts were significant (p>.05).

Time to contact with lead vehicle.

Comparison of autonomous conditions (fog vs. no fog).

The time to contact (TTC) the lead vehicle incorporates the speed and distance to lead vehicle at the start of the hazard. As before, participants was used as the random factor and screen manipulation (no-fog, with-fog) was the fixed factor. In line with the braking RT, speed and headway analyses, no significant difference was found between the no fog condition (4.4 sec) and the fog condition (4.3 sec), $t(245)$ =.035, p =.72, Marginal R^2 = .0005, Conditional R^2 = .0005.

Comparison of manual condition with autonomous condition (no fog).

Also as before, the second model used participants as the random factor and trial type (manual cones hazard, autonomous drive crash hazard no fog, autonomous drive cones hazard no fog, autonomous drive crash hazard with fog, autonomous drive cones hazard with fog) as the fixed factor. There was an overall significant effect of trial $t(207.3)$ =2.06, p =.04, Marginal R^2 = .08, Conditional R^2 = .09. However, once again post-hoc contrasts showed no significant difference in the comparison of interest, between the 'manual cones hazard' condition and the 'autonomous drive cones hazard no fog' condition (p>.05).

Results from remaining trial comparisons.

Again, for completeness I note that post-hoc contrasts showed a significant difference between the ‘autonomous drive cones hazard no fog’ condition (3.9 sec) and the ‘autonomous drive crash hazard with fog’ (3.5 sec; p =.03). However, this comparison is not meaningful because it confounds fog presence and hazard type. There was also a significant difference between the ‘autonomous drive cones with fog’ condition (5.2 sec) and the ‘autonomous drive crash hazard with fog’ (3.5 sec; p =.003). This indicates that, following the handover from the autonomous condition with fog, TTC was greater for an upcoming cone hazard than for an upcoming crash hazard.

Lane position analysis.

The distance in centimetres from the centre of the lane was calculated for all autonomous drive conditions for 60 seconds following the start of the takeover request. Smaller lane position values represent a vehicle position that is closer to the centre of the lane. For the model, participants was used as the random factor, with screen manipulation (no-fog, with-fog) and time bin (60 x 1 sec time bins) as the fixed factors. There was no overall effect of screen manipulation (no fog = 54.2 cm, fog = 52.7 cm), $t(287000)$ = 0.08, p =.91, Marginal R^2 = .17, Conditional R^2 = .26. The effect of time bin was significant for all times (p <.0001) except for the 0-1, 1-2, 2-3 sec time bins (p >.08). This result is unsurprising, because the autonomous driving system was still in control of the vehicle for the three seconds following the takeover request. The interaction between screen manipulation and time bin was significant at a large number of time points. For the sake of brevity, Table 6.1 below covers only the first ten seconds following the start of the takeover request. The results for the full 60 seconds can be seen in Appendix A.

Table 6.1. Mean distance in centimetres from the lane centre and standard deviation as a function of time bin (0-10 sec) and fog condition.

	No fog		With fog		
Time bin (sec)	Mean	SD	Mean	SD	Significance
0-1	1.2	1.7	1.2	1.6	$.97^{ns}$
1-2	1.7	2.0	1.8	2.0	$.99^{ns}$
2-3	0.9	1.0	0.9	1.0	$.96^{ns}$
3-4	3.6	5.6	3.5	5.8	<.0001***
4-5	27.6	25.1	23.9	25.0	<.0001***
5-6	64.3	43.9	53.4	38.0	<.0001***
6-7	69.3	43.4	62.1	36.7	<.0001***
7-8	63.3	41.2	64.5	40.0	$.36^{ns}$
8-9	61.4	43.5	64.4	42.6	.03*
9-10	62.1	44.4	59.1	38.9	.03*

(^{ns}= non-significant, **p*<.05, ***p*<.01, ****p*<.0001)

As can be seen in Table 6.1, there was no difference in lane positioning between the fog and no fog conditions in the first three seconds following the handover request. This is not surprising, because the automated system was controlling the lane positioning during that time. Over the following four seconds, drivers positioned the vehicle significantly further from the centre of the lane in the no fog (vs. fog) condition. However, it is important not to overstate the importance of these results, given that the maximum difference between the fog and no fog conditions amount to around 11cm, which is not particularly meaningful in the context of a normal driving task. The remaining three seconds for which data are shown in the table demonstrate a more varied pattern of results, without any systematic pattern of difference between the fog and no fog conditions.

Eye tracking analysis.

Fixation time back to the screen.

The first fixation time (FFT) back to the screen from the start of the takeback warning signal was calculated using the saccades package in R (Malsburg, von der, 2017). Any fixations of less than 100ms or more than 2000 ms were removed. Additionally, any RT to look back at the display over 5000ms were also removed. For the model, participants was used as the random factor and screen manipulation (no-fog, with-fog) was the fixed factor. There was a significant difference in FFT between the no fog condition (mean = 1989ms) and the fog condition (mean = 1566ms), *t(182.72)* =2.73, *p* =.007, Marginal R^2 = .02, Conditional R^2 = .24. See Figure 6.4.

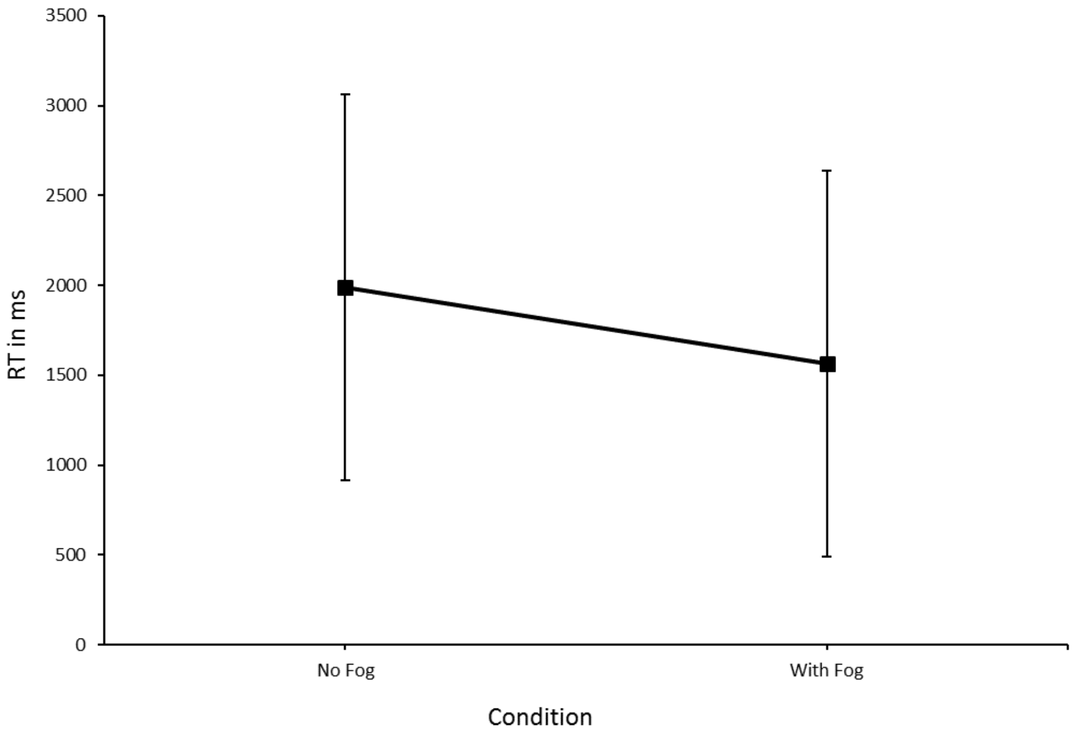

Figure 6.4. Time taken to make first fixation back to the screen after a takeover request in milliseconds (error bars represent 95% CI).

In line with the analysis of the time taken for participants' hands to return to the wheel, participants were quicker to return their eyes to the screen in the fog (vs. no fog) condition. Again, it seems likely that these results reflect an alerting effect in the fog condition, due to the abrupt removal of the fog causing a sudden visual change at the point of the handover request. This suggests that the presentation of an artificial stimulus creating a similar level of visual change at the point of the handover request (e.g. a large flashing stimulus on the windscreen) could improve reaction times to the request. However, as this was not the focus of this study, more research will be needed before firm conclusions concerning this possibility can be drawn.

First fixation duration.

The duration of the first fixation back to the screen was also compared between the fog and no fog conditions. Once again, any fixations of less than 100ms and more than 2000ms were removed. For the model participants was used as the random factor and screen manipulation (no-fog, with-fog) was the fixed factor. There was no significant difference between the no fog condition (358 ms) and the fog condition (328 ms), $t(196.56)$ =1.05, p =.31, Marginal R^2 = .005, Conditional R^2 = .008.

Percentage of road centre.

The percentage of road centre (PRC) was calculated using the eyetrackingR package (Dink & Ferguson, 2015) by finding the median fixation point for each individual participant and then constructing a square interest area of 6^{o} by 6^{o} visual angle. The number of gazes that fell in the interest area was then expressed as a proportion of the total number of gazes. The higher the proportion of gazes in the road centre, the more focused on task relevant information the participant is assumed to be. The PRC was only calculated for the comparison between fog and no fog conditions, as this was the main analysis of interest for this study. The basic data preparation was conducted for track-loss (i.e. when the eye tracker lost the eyes, through blinks, looking away or very low validity of gaze capture). There are no clear guidelines for what is acceptable tracker loss for this type of analysis. By comparison in an early word learning study of 18-48 month olds, a track loss threshold of 25% was used (Yurovsky & Frank, 2017). Because of the challenging nature of the eye tracking in this study (due to the fact that participants were encouraged to look away from the screen and down at their phones or tablets during periods of automation) I took a slightly more liberal threshold of 40% track loss on individual trials and 40% track loss for individual participants (i.e. if an individual trial for one participant had more than 40% track loss it was removed from the analysis and if an individual participant had more than 40% track loss across all trials they were removed). This resulted in 8 trials being removed and the data from 1 participant.

For the model, participants were used as the random factor and screen manipulation (no-fog, with-fog) and time bin (60 x 1 sec time bins) as the fixed factors. I used a logit adjusted proportion of PRC for the analysis but for clarity of interpretation I will report the proportion (%) of PRC. There was a significant effect of screen manipulation in the first second following the handover request signal (p =.04, no fog mean PRC= 55.6%, fog mean PRC= 36.9%, see figure 6.5). Thus, fixations during the first second were more dispersed following an automated period without visual scene information than following an automated period where visual scene information remained available. This finding could be seen as reflecting an increased need to examine the entire scene following the fog condition, due to the reduced situation awareness that is likely to have occurred during that condition (vs. the no fog condition). These results might at first appear difficult to reconcile with the mean values from the first fixation back to the screen analysis (no fog condition mean = 1989 ms, fog condition mean = 1566 ms), which suggested that many participants would not have returned their gaze to the screen in the first second following the handover signal. However, this difference is likely to be due to the way the two measures were analysed. In the PRC analysis, all fixations to the scene were analysed regardless of their duration i.e. very short (<100 ms were included) and long (>2000 ms) fixations were included in the proportion analysis, whereas in the FFT analysis very short and long fixations were excluded from analysis. This therefore means in the FFT only a first fixation greater than 100 ms would be counted as the first fixation, as previous research shows that fixations are very rarely less than 100 ms and often in the range of 200-400 ms (Salvucci & Goldberg, 2000) this seemed like a reasonable assumption to make. This therefore means that the amount of data included in each analysis was different, and the FFT analysis will have had a smaller amount of usable data compared to the PRC analysis.

The next significant difference between the fog and no fog conditions occurred 15-16 sec after the handover signal (p =.02, no fog mean = 74.7%, fog mean = 68.2%). This is in line with the previous finding, such that fixations were more dispersed in the fog (vs. no fog) conditions. And indeed, similar patterns of significant differences also arose at some subsequent time points, namely: 28-29 sec (*p* =.03, no fog mean = 75.4%, fog mean = 67.1%); 29-30 sec (*p* =.01, no fog mean = 77.6%, fog mean =68.1%); and 53-54 sec (*p* =.03, no fog mean = 85.2%, fog mean = 79.5%). These differences were not significant at any other time points (*p* >.06 for all comparisons). Overall, although significant differences between the fog and no fog conditions did not arise at all time points in the 60 seconds following the handover request, where there were significant differences they were always in the same direction, such that fixations were more dispersed following the fog (vs. no fog) condition. Taken together, these results suggest that drivers may have been taking in a small amount of visual information in the no fog condition, despite being engaged with their devices, such that they subsequently needed to spend less time scanning the scene very broadly when the automated system disengaged.

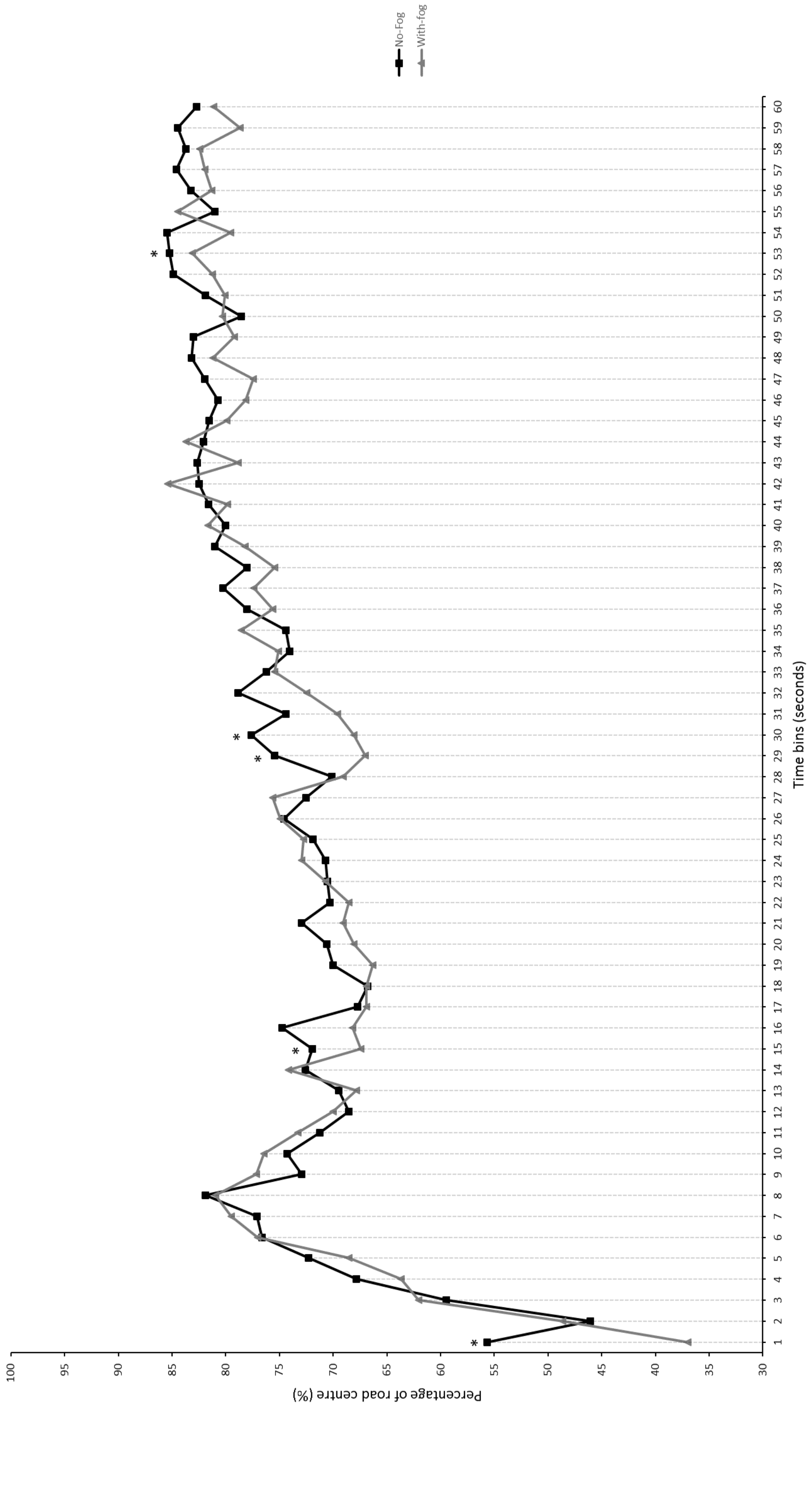

Figure 6.5. Percentage of road centre (%) for the two fog conditions. *$^*p<.05$*

Overall Discussion

In the present study, I examined the influence of providing visual information during a period of autonomous on both driving behaviour (lane position, headway, time to contact and speed) and eye movement behaviour (percentage of road centre, first fixation time and average fixation duration). Overall, my two goals were to examine the effect of providing or occluding visual information during a period of automated driving on takeover performance and how responding to a hazard might differ following automated driving as compared with manual driving.

One of the most consistent findings was that takeover responses were quicker when the period of automated driving included heavy fog (and thus visual information from the driving scene was occluded) compared with when the automated period did not include fog. This finding was the same whether takeover response was measured in terms of the time taken to make a fixation back to the screen or the time taken to return the hands to the wheel. As discussed above, it is likely that the abrupt removal of the fog caused a visual transient which captured visual attention. However, it is also possible that drivers were aware of their reduced access to visual information during the fog condition and therefore prioritised the takeover process more highly in fog (vs. no fog) conditions. Future work could examine this possibility directly by removing the fog more gradually, so as to avoid creating such a large visual change.

It is also interesting to note the large standard deviations in both of these takeover measures (time taken to return hands to the wheel and to make a fixation back to the screen). This is in line with previous research (e.g. Eriksson and Stanton, 2017) identifying large individual differences in the time taken to regain control. It is also likely that the variation in the tasks that the participants were engaged with could have contributed to these results. Participants were given free choice of the task to carry out on their device during automated driving, so it is possible that those who were doing a task that interested them and engaged them more may have taken longer to fixate away from the device and put the device down than those who had chosen a less engaging task. Future research could examine this possibility by controlling the task that participants carry out during the period of automated driving, although this would lose some of the ecological validity that was present in the current study, because in real world operations drivers will have control over the additional tasks that they perform. These results have implications for the takeover request times used by the automobile industry or policy makers. It might at first seem prudent that they should endeavour to use transition times that are as long as possible, so that they do not exclude users with reaction times that are longer than average. However, this idea may not be feasibly practical, particularly because the need for the human driver to take over will not always be detected with a great deal of warning. Nevertheless, it might be possible through settings present in the autonomous system to tailor the takeover request time to individual drivers, and this could be an interesting area for future research.

The PRC analysis which is argued to be a sensitive indication of visual distraction showed that in the first second after a takeover request visual attention was more dispersed in the with-fog to the no-fog condition. This shows that when visual information has been reduced during automated driving, a driver is more likely to look holistically at the visual scene during the subsequent takeover process, rather than focussing on a central point where most of the important task relevant information is available. This is indicative of a driver in the with-fog condition reorienting themselves back into the visual scene and the task of driving by widely scanning the environment. Importantly, this reorientation seems to be quick and resolves after 1 second. The observed difference in PRC between the two conditions at 15-16 secs, 28-30 sec and 53-54 sec were always in the same direction, such that scanning was more dispersed following fog (vs. no fog) in the period of autonomous driving. It is possible that this may indicate a similar reorientation process occurring from time to time across the 60 seconds. However, this seems unlikely because visual attention is likely to resolve to a similar level fairly quickly following the handover, so it seems implausible that visual occlusion during the automated period would still be having genuine effects. If that were the case, I would expect to see more differences earlier after the takeover request, but these were not apparent.

The driving behaviour measures show that the braking RT, speed, headway, and TTC at the start of the driving hazard do not significantly differ between the two occlusion conditions. This indicates that in the approximately 3-4 seconds after the driver takes back control of the vehicle they regain a similar amount of vehicular control. However, the lane position analysis shows that at the takeover point (after the 3 second takeover warning) where the vehicle is back under manual driver control, there is a significant difference in lane position between the conditions. In the initial 4 seconds (between 3-7 secs in table 6.1) drivers in the no-fog condition on average deviated further from the centre of lane than those in the with-fog condition. As argued in relation to the takeover reaction time results, it is possible that drivers in the with-fog condition were aware of their reduced access to visual information and therefore prioritised the takeover process more highly than drivers in the no-fog condition, which could explain why drivers in the with-fog condition appeared to exhibit slightly better control of lane position. However, these differences between the two conditions are relatively small, in the range of 11 centimetres at the most, and differences were not observed in any of the other measures of vehicle control, so it is important not to overstate the implications of these results.

Responding to a hazard was observed to differ following automated driving as compared with manual driving in the current study. Specifically, participants tended to drive more slowly in the autonomous condition at the start of the driving hazard, perhaps suggesting that drivers are more cautious in the period following the handover and therefore adopt lower speeds. However, this cautious driving behaviour was not mirrored in the headway and TTC analysis. Additionally, braking in response to a developing hazard was no slower following an autonomous handover than following a period of manual driving. Taken all together these results show that after a period of automation and a short 3-4 sec duration after takeover before the hazard starts, drivers in the autonomous conditions show very few differences in driving performance compared to manual drivers. This potentially indicates that drivers are able to quickly and effectively reengage with manual driving following a period of autonomous driving. The patterns of results is similar to those found by Young and Stanton (2007) who found no differences in braking response time to a hazard between autonomous and manual driving. But this contrasts with the results found by Radlmayr, Gold, Lorenz, Farid, and Bengler (2014) demonstrating slower responses to hazards after a period of automation compared to a manual driving condition. Additionally, my results show similar takeover performance as found by Mok et al. (2017), and a similar pattern of visual attention as found by Louw et al. (2016) in their studies. However, the results of my study need to be qualified in the sense that the driving environment used was benign, i.e. swarm vehicles were not decelerating rapidly and there were no substantial speed differentials between the ego and swarm vehicle. Although drivers were able to maintain the "status-quo" of vehicular control, they might not be ready for severe avoidance manoeuvres after a period of autonomous driving.

Although the handover requests in the present study occurred at a similar speed in all conditions (60-70 mph), it is important to note that the speed of the vehicle at the moment of handover can affect the speed and accuracy with which the driver is able to regain control of the vehicle. The Venturer Trial 1 assessed this using two methods: a driving simulator and a road driven vehicle (Morgan & Alford, 2017). In both methods, the key experimental variables were takeover (time taken to reengage with vehicle controls) and handover (time taken to regain a baseline/normal level of driving) behaviour. Performance was measured when switching frequently between automated and manual driving within urban settings during relatively short (4 -8 mins) driving scenarios. The IV in this study was speed (20, 30, 40, and 50 mph) and the DVs included lateral acceleration, lane position, RT to regain manual controls after a handover request and average speed. One of the unique aspects of this study's methodology is that it allows for direct comparisons between the behavioural results from the simulator and real-world driving. 31 participants took part in the simulator study and 27 participants took part in the real-world road driven vehicle study. In the simulator study, the takeover time analysis revealed that takeover time was significantly higher in the 20-mph condition than the 30 mph ($p < .01$), 40 mph ($p < .05$), and 50 mph ($p < .001$) conditions. More specifically, the minimum amount of time to take back the simulator controls after the handover request was 1.99 seconds in the 50-mph condition rising to 2.47 seconds in the 20-mph condition. Takeover time in between the 30, 40, and 50 mph conditions did not differ statistically (*ps*> .05). In the next set of analyses, a baseline manual condition was compared with the handover phase (over 55 secs) for the IV and several different DVs. The standard deviation lateral position (SDLP), measured in metres from the centre of the lane, is an index of a drivers 'weaving' in a lane, and is a stable measure of driving performance with high test–retest reliability (Verster & Roth, 2011). Analysis of this measure revealed a difference between the handover period and manual baseline only in the 50-mph condition, such that there was more weaving during handover (vs. baseline). This indicates that participants were less able to maintain a stable lane position during handover at this higher speed. The steering input results (how many times the driver moves the steering wheel) showed that in all but the 30 mph condition participants tended to move the steering wheel more often during handover than in the baseline condition suggesting less controlled steering during handover. In summary, it took approximately 2 seconds to take back manual vehicle controls following a handover request in the 30-50 mph conditions, with a slight slowing in the 20-mph condition. At some speeds, there was also lower control over the vehicle (steering input) and less appropriate positioning within lane during handover (vs. the baseline condition).

These results make an interesting comparison with my study, in particular in relation to the handover times. More specifically, whereas Morgan & Alford (2017) found reaction times of around 2 seconds for taking back manual control following the handover request, I found reaction times of around 4 seconds. This is likely to reflect the fact that participants in the current study were engaged in a secondary task (using their phone/tablet) and this is likely to have slowed down their motor RTs.

In the real-world driving task, Morgan and Alford (2017) used the IV of driving mode (manual driving and driving after handover) and a range of DVs (including average speed, SD of average speed and SDLP). The speed of the autonomous vehicle was limited to a maximum of 20 mph. The handover time was 1.73 sec (SD=.81) at an average speed of 14-17 mph. This handover time is shorter in duration than the comparable speed (20mph) in the simulator condition, perhaps suggesting that people were not prioritising the takeover process to the same extent in the simulator as they did in real driving conditions. The driving speed and SD speed were compared to the baseline, over 0-1 second, 0-5 seconds and 0-10 seconds after handover. The manual baseline speed recorded at the same location as the assessed handovers was close to 20 mph. All other handover speeds were lower than at baseline and decreased with longer post- handover period of assessment (baseline - 20 mph; 1 sec after handover - 17 mph; 5 secs after handover - 15 mph; 10 secs after handover - 14 mph). This result suggests that drivers slow down after control is transferred back to them. The SD speed showed a reduction in speed variation at 1 sec after handover in comparison to baseline, speed variation then increased at 5 secs after handover and then further at 10 secs after handover.

In summary, the results of both the simulator and real-world studies show that the handover time is impacted by the autonomous vehicle's speed at the point of handover, with longer handover times in the lower speed condition. This potentially indicates that the drivers believe that the risk of the vehicle swerving or changing course is lower at slower speeds and therefore treat the handover process with less urgency than they do at higher speeds. As for the SDLP results, these show that vehicle control and lane positioning are easier to manage at lower speeds and only in the fastest speed condition (50 mph) was there a significant difference from the baseline condition. It is therefore possible that different results would have been found from the current study had lower handover speeds been used, and this could be an interesting area for future research.

For reasons of practicality, the periods of automation that I used in the current study were fairly brief (i.e. 7-8 minutes) whereas takeover requests in future AVs are likely to occur less frequently. To examine if there are any differences between short or long periods of automation, Feldhütter, Gold, Schneider and Bengler (2017) compared driving performance, eye movement patterns and RTs during a takeover after 5 or 20 minutes of automated driving in a driving simulator. They also examined the effect of two levels of task load (no additional task, and visual pattern matching task) on takeover performance. There was a significant difference in RT – defined as the time that the participants need for directing the first gaze away from the secondary task after the takeover request – between the 5-minute automation condition (.55 sec) and the 20-minute condition (.64 sec). However, there was no difference between the two conditions in take over time – defined as the time that the participants need to start a manoeuvre as a reaction to the takeover request. Thus overall, there is a slight slowing of the first fixation back to the central visual scene following a longer period of automation, but for all other measures there is no difference of the length of automated driving period that occurs before the handover. This potentially indicates that the results of the current study should be broadly comparable to studies that use longer periods of automation.

One of the design features that have been explored by vehicle manufactures to help drivers reengage with driving after a period of autonomous driving is to periodically prompt the driver to take over control the vehicle. Although this technique was not used in the current study, there is evidence to suggest that it can improve takeover performance. For example, in a simulator and eye tracking study, Merat, Jamson, Lai, Daly and Carsten (2014) examined drivers' ability to resume control from a highly AV in three conditions: one in which automation was switched off and manual control was required at regular 6 minute intervals (fixed); a second condition where the transition back to manual driving was based on the duration of time drivers were looking away from the screen (variable); and a third manual drive condition that was used as a baseline. 46 participants completed the three driving conditions in a high fidelity driving simulator. Merat and colleagues measured percentage of road centre (PRC) and SDLP over the 60 seconds following the handover. The results showed that PRC as measured across the entire 60 seconds following the handover did not vary significantly between the fixed and variable conditions. However, differences were revealed when the PRC was broken down into 12 five-second time chunks. In the variable condition (vs. the fixed condition), the PRC was generally lower during the first minute of manual control. Specifically, in the fixed condition, the PRC stayed stable over the first 10 seconds following the handover, but in the variable condition the PRC started off lower than the fixed condition (indicating more of a spread of fixations) but quickly increased over the first ten seconds to a level that was comparable with that in the fixed condition. Additionally, PRC in the variable condition significantly decreased between 30-50 secs whereas PRC in the fixed condition remained relatively stable over this period. This result would indicate that in the variable condition the driver's visual attention is more diffuse than in the fixed condition after the initial takeover request, and after a short period of stabilisation again dips before recovering to a similar level as in the fixed condition. SDLP was not significantly different between the two conditions when collapsed across the entire 60 seconds following the handover request. However, for both conditions, SDLP was much lower during the first 10 secs (approximately .05m) than during the remaining 50 seconds, during which time SDLP stabilised to between 0.1- 0.2m. This indicates that drivers are likely to spend some time orientating themselves within the driving environment before they start to make changes to their lateral position. In summary, these results importantly show that there is a quantifiable difference between a fixed or variable schedule of takeover requests in patterns of some aspects of driving behaviour and visual attention. The implications of the results are that a fixed schedule may work better at reengaging drivers with manual driving rather than a schedule in which the handover is only requested when the driver's attention drifts. However, this may not be a practical method for keeping drivers in the loop and future research is needed.

The current study used an auditory takeover request signal presented from speakers within the computer monitors. Indeed, in almost all studies in this area takeover requests are presented either visually or auditorily and unidirectionally, with little research identifying if these modalities are the most optimal ways to present this information. Petermeijer, Bazilinskyy, Bengler and de Winter (2017) conducted a simulator-based study to investigate the effects of takeover request modality and directionality on drivers' steering behaviour when facing a collision scenario. 24 drivers drove an AV which could issue takeover requests in different modalities: auditory, vibrotactile, and auditory-vibrotactile. The auditory stimuli were presented via speakers in the simulator and the vibrotactile stimuli via a seat that included tactors on both sides. Regardless of modality, takeover requests could be presented from the left, from the right, or from both the left and the right. The takeover request warned the drivers about a stationary vehicle that had to be avoided by changing lane. The researchers measured the hands to the wheel RT, braking RT and lane change RT. Interestingly, the auditory-vibrotactile (1.5 sec) requests induced significantly faster hands to the wheel RT than the either the auditory (1.69 sec) or vibrotactile requests (1.8 sec). There was no difference in hands to the wheel RT time between the unimodal cues. Additionally, the braking RT and lane change RT did not significantly differ between either of the three conditions. Finally, the direction and congruency of the stimuli (i.e. if the stimuli were presented from the left and the driver needed to move left) did not influence any of the RT measures. In summary, the results of this study show that a multimodal auditory-vibrotactile cue can speed up the RT for driving returning their hands to the wheel and start regaining control of a vehicle compared to unimodal cues. There seem to be no benefits of directional auditory and vibrotactile stimuli in prompting a driver to change lane if required to by the autonomous system. However, it is possible that the saliency of the cues and the relative novelty of the presentation of the cues may have made it more difficult for the drivers to utilise the cues effectively.

Participants in the current study had only a brief practice period in which to become familiar with the handover process. Although this was necessary from a practical standpoint, it may mean that handover performance was underestimated in the present study, because such performance has been shown to improve with practice. For example, Hergeth, Lorenz and Krems (2017) examined the effects of prior familiarisation with takeover requests during level 3 automation on drivers' takeover performance. Using a driving simulator, takeover performance was examined across four familiarisation groups (no familiarisation, description, experience, description-and-experience). Takeover time was significantly longer in the no familiarisation (3.5 secs), and experience conditions (2.9 secs) than in the description-and-experience condition (2.3 secs). There was no difference between the description condition (2.5 sec) and the other conditions. Thus, prior familiarisation with takeover requests can reduce the RT to take over control of the vehicle after a system request to do so, particularly if the familiarisation includes both description and experience. However, there were no effects of familiarisation condition on the measure of time to collision with the lead vehicle at the takeover request (i.e. the time after which two objects would collide if they maintained their present speed and trajectories).

Finally, there is evidence that takeover processes can be affected by the level of trust that drivers have in the autonomous system. For example, Payre, Cestac and Delhomme (2016) conducted a study to assess the impacts of practice, trust, and interaction on manual control recovery when resuming control of a level 4 AV in a driving simulator. 69 participants were split into two groups: one group had limited practice of resuming control of the vehicle after a period of automation; the second group received additional practice on this task. All participants completed a questionnaire before and after the driving session that probed their trust in the autonomous driving system. There was no significant difference in the trust of the system before or after experiencing the autonomous driving system. Thus, in this study at least, the level of trust in the system was not influenced by exposure to the driving system. However, participants that had a higher trust score showed longer RTs to resume control of the vehicle than participants with a lower trust score, and these effects were not modulated by the amount of practice available. These findings suggest that a driver with a higher level of trust in the system might also exhibit a higher level of complacency toward the system, such that they are more likely to assume that the system is capable of dealing with complex situations without manual input, and they therefore allocate less attention from the driving task. By contrast, drivers with low levels of trust might be less complacent, constantly monitoring what the system is doing and the driving environment, and therefore are more likely to be primed to take over vehicular control at short notice. This theory is broadly congruent with other work on automation (not driving-specific) that shows that complacency in autonomous systems due to lack of understanding of the system's boundaries and limitations can influence how people respond towards them (Bainbridge, 1983; Hoc, Young, & Blosseville, 2009). Thus, although I did not measure or manipulate participants' levels of trust in the autonomous system in the current study, it is possible that pre-existing differences in their levels of knowledge and acceptance of automation may have influenced their takeover performance, and this may have contributed to the large individual variability that I observed. Given the more general importance of people's trust in automation for the eventual acceptance of driverless technology, my final empirical study focused on measuring current attitudes to autonomous vehicles.

Chapter 7 – Attitudes to Autonomous Vehicles Survey

Introduction

"When a train goes through a tunnel and it gets dark, you don't throw away the ticket and jump off. You sit still and trust the engineer". Corrie Ten Boom

Whereas the previous chapters have looked at dual task performance, situation awareness and visual attention, I am now turning to examine people's attitudes to automated vehicles. This is important because the route to full automation is likely to involve significant challenges, with public attitudes playing an important role in determining the level of success with which the technology is introduced. Trust in automation and in particular trust in AVs is important for the eventual uptake and success of AVs. There already exists a lot of research on this topic, as outlined in the following sections, but the survey reported in the current chapter includes an additional experimental manipulation, designed to investigate whether providing up-to-date information on the future impact of AVs influences people's opinions of AVs.

General Attitudes Towards AVs

There already exists a rapidly-growing body of surveys on public attitudes to vehicle automation. For example, Gkartzonikas & Gkritza (2017) identified 25 such surveys published since 2012. In general, the findings remain unclear concerning the public's overall level of interest in using or buying the technology (e.g. Cavoli, Phillips, Cohen, & Jones, 2017). However, there is fairly good agreement regarding the most frequently identified concerns:

1) Safety and reliability of the system: safe functioning is often rated as people's top priority when judging the desirability of autonomous vehicles (e.g. Bansal & Kockelman, 2016; Schoettle & Sivak, 2014)
2) Security of the software: respondents typically raise the potential for people to hack into vehicle control systems as another serious concern (e.g. Kyriakidis, Happee, & de Winter, 2015)
3) Cost: there is an assumption that autonomous vehicles will increase the cost of vehicle ownership (e.g. Bansal & Kockelman, 2016; Howard & Dai, 2014)

4) Liability: several surveys identify concerns around the legal issues associated with use of the technology (e.g. Howard & Dai, 2014; Kyriakidis et al., 2015).

Given this level of agreement concerning the most frequently-raised concerns around the introduction of AVs, the survey used in this chapter did not ask participants directly about their concerns, but instead focused on their trust and acceptance of the technology.

In order to understand what factors influence trust in AVs, Choi and Ji (2015) developed a structural equation model of trust and acceptance of AVs. A survey of 552 Korean drivers was conducted, focusing on drivers' acceptance of new technology, how much they trust the technology and the perceived risk of the new technology. They identified 10 constructs (such as perceived usefulness, perceived ease of use, trust, perceived risk, system transparency, technical competence and situation management) to be used in the model and tested the strength of the hypothesised relationships in the model and the robustness of the model in predicting behavioural intention to use AVs. Perceived usefulness and trust were the major determinants of intention to use AVs, with some of the major determinants of trust levels being identified as: system transparency, technical competence, and situation management. Trust also had a negative effect on perceived risk i.e. as trust in AVs increased their perceived related risk decreased. There is a potential limitation to this study, being that they had a relatively high proportion of younger (<35 yrs.) respondents (who may be more open to technological developments) as well as more male than female respondents. The potential problem with this gender imbalance is that there may be a different pattern of responses between the genders, as I will outline in the following section.

Gender Differences in Trust and Acceptance Of AVs

There are notable differences between the sexes in terms of their willingness to use AVs, with males usually reporting higher tendencies to use AVs than females (Ernst and Young, 2013; Kyriakidis et al., 2015). Hohenberger, Spörrle, and Welpe (2016) investigated the possibility that affective reactions (the emotional or mood response to a situation, e.g. pleasure or anxiety) might be able to explain behavioural intentions and responses towards AVs, and that these effects might vary depending on sex and age. Using an internet-based survey methodology they sampled 1603 German participants asking five questions in relation to a description of a level 2-4 SAE vehicle ("How frightening would such a car be for you?, How much pleasure would driving such a car provide for you?, Were you willing to use such a car today?, Age, and gender"). Using mediation analysis, they found that affective responses towards AVs (e.g. anxiety or pleasure, measured using a Likert scale in relation to the questions above) explained the effect of biological sex on willingness to use AVs, such that the female group which exhibited higher levels of anxiety were less likely to respond that AVs would be pleasurable or they would be willing to use them. Males showed lower levels of anxiety and higher levels of willingness to use and pleasurableness from AVs. Additionally, age was an important factor in determining the level of anxiety around AVs across all participants, with the level of anxiety decreasing as age increased. The results of this study indicate two clear messages: one, the uptake and willingness to use AVs is influenced by gender, with males reporting lower levels of anxiety than females; and two, increasing age moderates the effect of anxiety in the willingness to use AVs, as older people feel less anxious than younger experienced drivers with AVs. Although, it was beyond the scope of the current study to attempt to measure anxiety, the effects of gender and age on responding will be considered.

The Effect of Providing Information About AVs on People's Trust and Acceptance of Them

Two of the main factors than might affect drivers' trust of automation are: 1) experience of automation; and 2) a better understanding about the pros and cons of automation. The former is harder to assess at present due to the relatively low prevalence of higher level (2/3 SAE) vehicle automation (although I did ask participants to report their prior exposure to AVs in the current survey). By contrast, the effects of providing information are more amenable to experimental study. Indeed, a small number of studies have assessed whether providing information to survey respondents can influence their trust in automation. For example, Souders and Charness (2016) provided older participants (55+) with information regarding AVs and different ADAS systems, and measured their trust in AVs. A sample of 459 took part in the survey with 188 receiving the additional information and 271 not receiving the additional information. The results showed that there was no difference between the control and information group for any of the demographics collected, highlighting that the groups were homogenous. The main survey results showed that for the combined data for both groups, the willingness to use AVs, the perceived benefits and familiarity with ADAS was all positively correlated with their trust in the system i.e. as trust increased so did the acceptance of the technology. The effect of providing information had little significant effect on the overall trust in automation, but did influence the level of concern about issues involving AVs and their safety. More specifically, those who received the information sheet reported less concern about AVs than those who did not. In conclusion, the results of this study show that providing information about AVs and ADAS systems to older drivers does have a small effect on their concern about AVs. However, the generalisability of the results is potentially limited given that the sample surveyed were all older drivers, who may not have views that are representative of the population as a whole (as demonstrated, for example, by Hohenberger, Spörrle, and Welpe, 2016, described above). One of the strengths of my survey is that I imposed no such restrictions, which should render the results more generalisable to the wider population.

Howard and Dai (2014) surveyed a wider age and gender range than Souders and Charness (2016) and also presented additional information to participants. They surveyed 107 people at an American science museum, and showed all participants a short 10-minute video that explained what AVs are and how the technology works. The results showed that participants found increased safety, the ability to multitask, and convenience to be the most useful features of AVs. Additionally, the results revealed that participants were interested in AVs as an improvement to their lives, and valued personal comforts and convenience higher than societal benefits (like environmental friendliness and reduced travel times). The aspect that was most concerning for most participants concerned who would be liable if there were a crash, the cost of AVs and control of the vehicle.

Finally, the males surveyed were more likely to be concerned with liability and less likely to be concerned with control than the females. Although it is possible that the presentation of information about AVs (via the video) may have influenced the responses of respondents, with no control group it is not possible to assess the effect.

The purpose of this study was first to gauge peoples' general views on AVs, and whether they trusted and accepted AVs on the road in the UK. Second, I wanted to examine if presenting information about the potential pros and cons about the effect of AVs on business, transport, society and individuals would influence their views of AVs. This study provides a novel way of examining the effect of information on people's perception of AVs by using a control group (with no information) and an experimental group (with information) to provide a direct assessment of the effects of providing new information about AVs to the public.

Method

An online survey examining people's trust and acceptance of AVs was conducted from December 2016 to February 2017 using Qualtrics. The majority of questionnaire items used in the survey were adapted from previous studies of general driving attitudes, trust and acceptance, and specific attitudes to AVs (AAA, 2016; Garcia, Kreutzer, Badillo-Urquiola, & Mouloua, 2015; Tennant et al., 2015). Additionally, I used an adapted version of the 12 question 'trust between people and automation' scale (Jian, Bisantz, & Drury, 2010). This is 7-point Likert scale made up of statements that evaluate trust between people and automated systems. I adapted the scale by changing the focus of the question from "system" to "a driverless vehicle(s)". For example, "I am suspicious of the system's intent, action or outputs" was adapted to "I am suspicious of a driverless vehicle's intent, action or outputs". Additionally, a range of demographic information was also collected, covering driving habits, driving experience, experience of driving aids, occupation and education. The full text of the survey, along with responses (including all freeform responses to an open section at the end asking for "do you have any additional comments?") is included in Appendix B and C.

As previous research has indicated that a lack of information may cause people to adopt a more cautious attitude to AVs, I also provided half of the respondents with up-to-date information on AVs (see information sheet in Appendix D) before they completed the main survey questions (i.e. after they had completed the demographic questions). Two comprehension questions were included for this group ("In the previous article what was the percentage of road traffic collisions thought to be caused by human error?" and "What was fatality rate of manual driving in the UK?") with the purpose of screening out participants who did not properly comprehend the additional information. The entire survey took between 20-30 minutes to complete.

An internet-based approach was used with the aim of recruiting a wide and diverse demographic. However, I specifically targeted members of public who had signed up to take part in the GATEway project. The GATEway (Greenwich Automated Transport Environment) project is an £8m research project, led by TRL, to understand and overcome the technical, legal and societal challenges of implementing automated vehicles in an urban environment. Taking place in the Royal Borough of Greenwich, the project is trialling and validating a series of different use cases for automated vehicles, including driverless shuttles and automated urban deliveries. The work aims to inspire and engage the public with the potential of automated transport technology and the project has recruited many supporters and those interested in trialling the technology (GATEway, 2017). The questionnaire was also disseminated through Twitter, email and the Royal Holloway, University of London intranet. The recruitment process resulted in 299 replies from potential respondents. Only surveys that were at least 85% complete and with at least one of the comprehension questions correct (if relevant) were included for further analysis, resulting in a final sample of 233 respondents. Finally, it should be noted that online surveys, by their nature, do result in the exclusion of individuals without computer or internet access.

Results and Preliminary Discussion

148 of the 233 respondents were recruited via TRL's Gateway project, having expressed an interest in taking part in research and trials relating to vehicle

automation. The remaining 85 people (henceforth referred to as the 'internet group') indicated that they had been recruited via other means (primarily email and social media). The GATEway group (mean age 41) were significantly older than the internet group (mean age 35; $t\ (231) = 3.38$, $p = .001$). The GATEway group also had a significantly larger proportion of male respondents (79%) than the internet group (60%; $\chi 2\ (1, N = 229) = 9.24$, $p = .002$). Because both age and gender are known to affect attitudes to autonomous vehicles (e.g. Cavoli et al., 2017; Hohenberger, Spörrle & Welpe, 2016), and in addition because the GATEway group were recruited based on an already-stated interest in driverless vehicles, the results are compared separately for the GATEway and internet groups on the two central questions ("Driverless cars are a good idea" and "I can trust a driverless vehicle") in acknowledgement of the fact that the GATEway group may be less representative of mainstream attitudes than the internet group.

Overall Attitudes to Autonomous Vehicles

Overall attitudes to autonomous vehicles can be gauged from responses to the two general statements "Driverless cars are a good idea" and "I can trust a driverless vehicle". The sections below examine the responses to these two questions.

"Driverless cars are a good idea".

The results from all participants shows that 81% replied 'agree' or 'agree strongly' and 1% replied 'disagree' or disagree strongly' with the statement above. Figure 7.1 shows responses from all participants to the statement 'driverless cars are a good idea'. In order to test for a difference between the GATEway and internet groups, answers were categorised as "agree" (comprising both the "agree" and "strongly agree" categories), "neutral" (comprising the "somewhat agree", "neither agree nor disagree" and "somewhat disagree" categories) or "disagree" (comprising the "disagree" and "strongly disagree" categories). A Fisher's exact test revealed a significant difference in the patterns of responding across the two groups ($p = .04$).

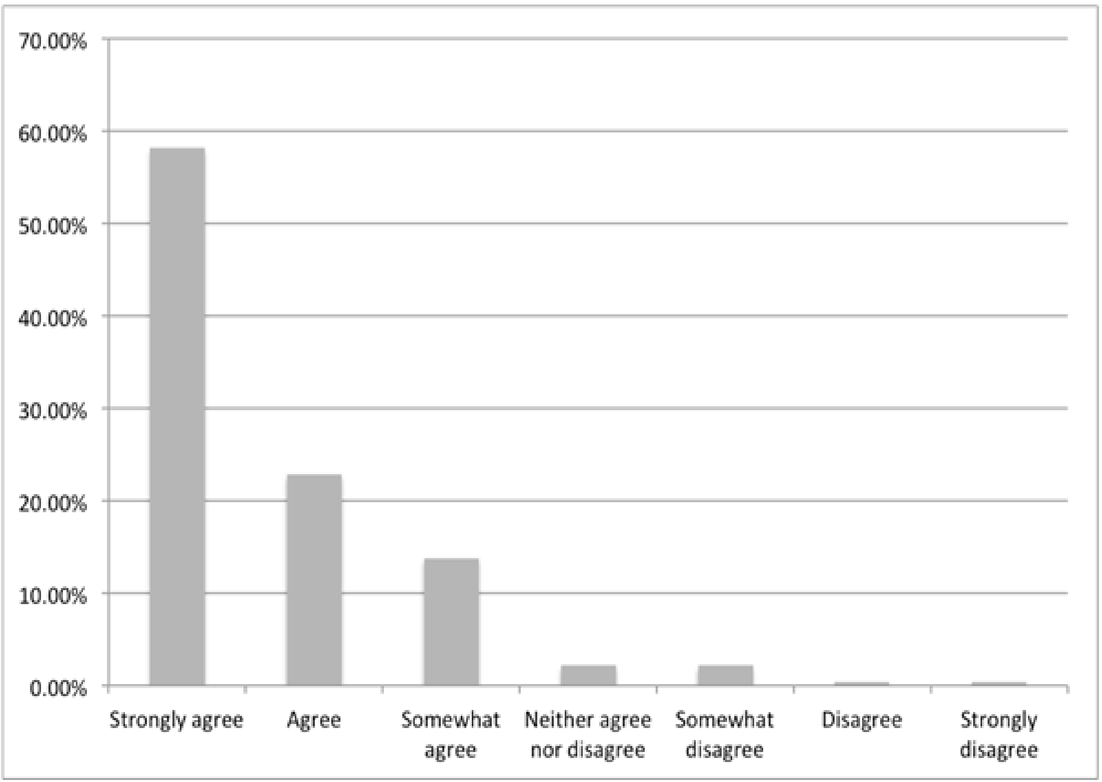

Figure 7.1. Responses from all participants to the statement “Driverless cars are a good idea”. Reprinted from Attitudes to autonomous vehicles, by Hyde, Dalton and Stevens 2017, Transport Research Laboratory (TRL). ISBN: 978-1-910377-91-8. Reprinted with permission.

Figure 7.2 therefore shows responses separately for the GATEway and internet groups. Note that, despite being less positive than the GATEway group, the internet group were nevertheless highly positive overall: 74% replied “agree” or “agree strongly” and 1% replied “disagree” or “disagree strongly”. In line with previous research (e.g. Hohenberger, Spörrle & Welpe, 2016) a Fisher’s exact test using the same response categories identified a significant difference between the response patterns of male and female participants ($p < .001$). Men were more likely to agree with the target statement than women (89% vs. 62%) and less likely than women to give a neutral response (11% vs. 36%). I did not undertake detailed qualitative analysis of participants’ replies to the freeform response section at the end of the survey (in which respondents were asked “Do you have any additional comments?”) However, some examples of these comments are provided throughout the report where they relate to key themes. Two examples of the most positive comments are provided below, taken verbatim from within the questionnaires:

> *"Bring them on!"* - respondent comment
>
> *"I work in the railway industry where digital signalling has meant that driverless trains are the norm for metro railways. I have seen first hand the improvements in safety the technology has delivered therefore I have no doubt it will deliver the same results for road transport. I can see the enormous benefit this technology will have for elderly and disabled people who no longer can drive and find it difficult to use public transport/ walking or cycling as mobility options."* - respondent comment

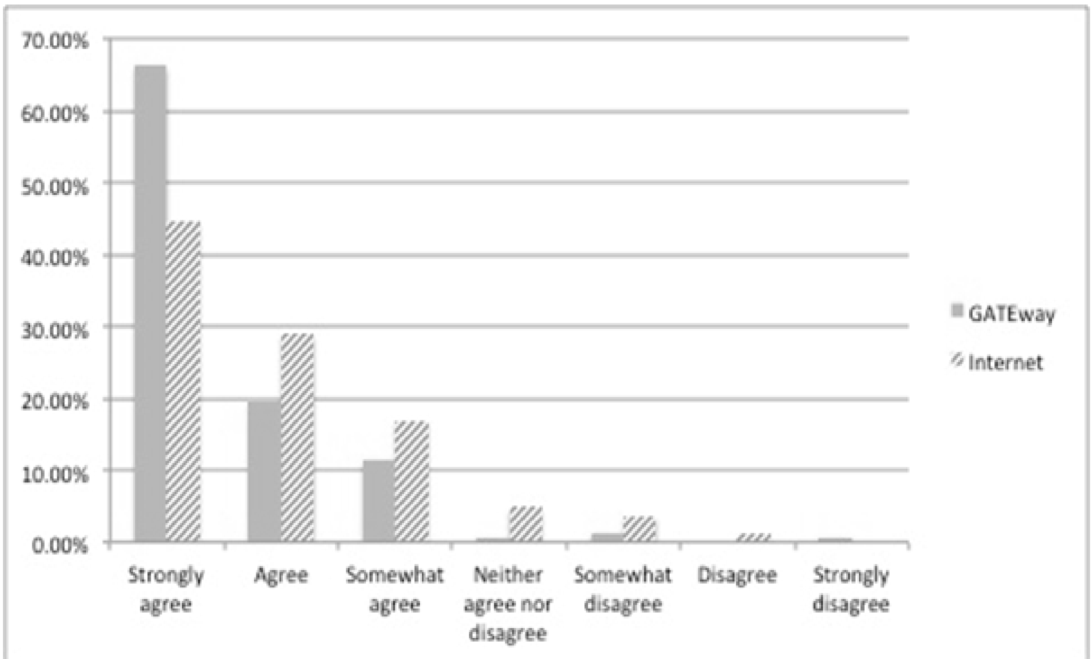

Figure 7.2. Responses to the statement "Driverless cars are a good idea" separated by group. Reprinted from *Attitudes to autonomous vehicles,* by Hyde, Dalton and Stevens, 2017, Transport Research Laboratory (TRL). *ISBN: 978-1-910377-91-8*. Reprinted with permission.

'I can trust a driverless vehicle'.

The results from all participants on this item demonstrated that 55% replied "agree" or "agree strongly", 23% replied "somewhat agree" and 3% replied "disagree" or disagree strongly'. Figure 7.3 shows responses from all participants to the statement "I can trust a driverless vehicle". Again, in order to test for a difference between the GATEway and internet groups, answers were categorised as described for the previous question. A Fisher's exact test again revealed a significant difference in the patterns of responding across the two groups (*p* =.004). Figure 7.4 therefore shows responses separately for the two groups. Again, although the internet group was in general less positive, their responses showed a similar pattern overall. A Fisher's exact test using the same response categories identified no significant difference between the response patterns to this statement of male and female participants (*p* =.13).

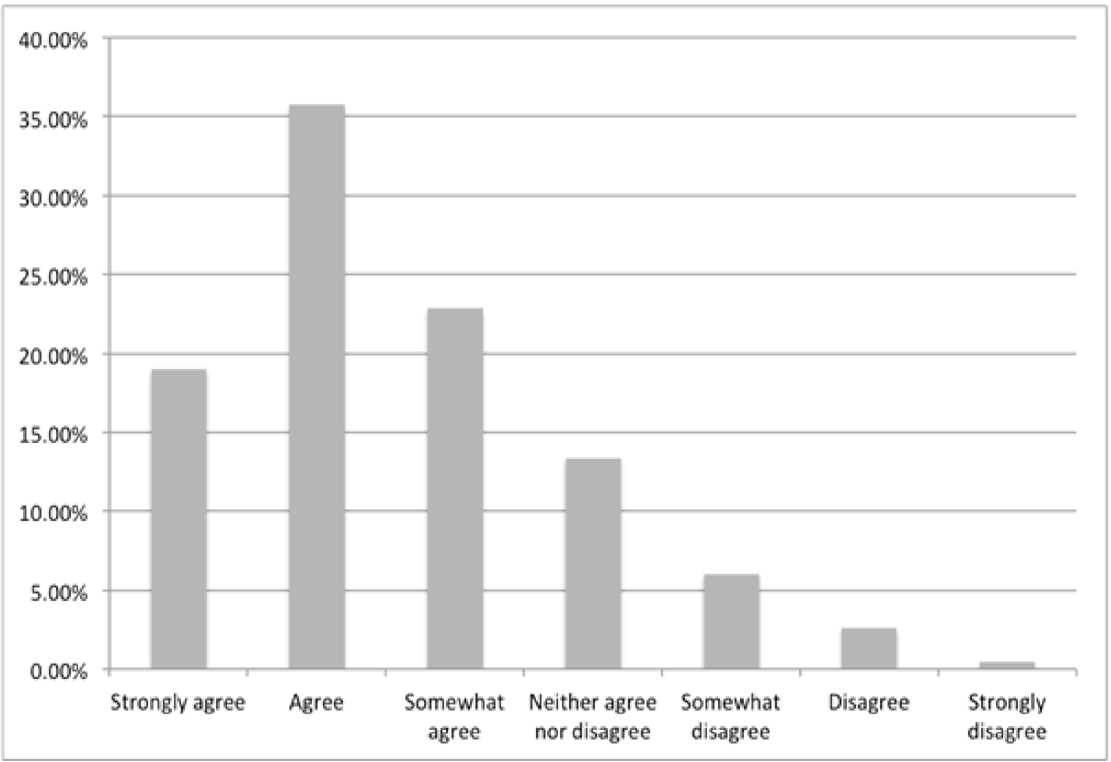

Figure 7.3. Responses from all participants to the statement "I can trust a driverless vehicle". Reprinted from *Attitudes to autonomous vehicles,* by Hyde, Dalton and Stevens, 2017, Transport Research Laboratory (TRL). *ISBN: 978-1-910377-91-8*. Reprinted with permission.

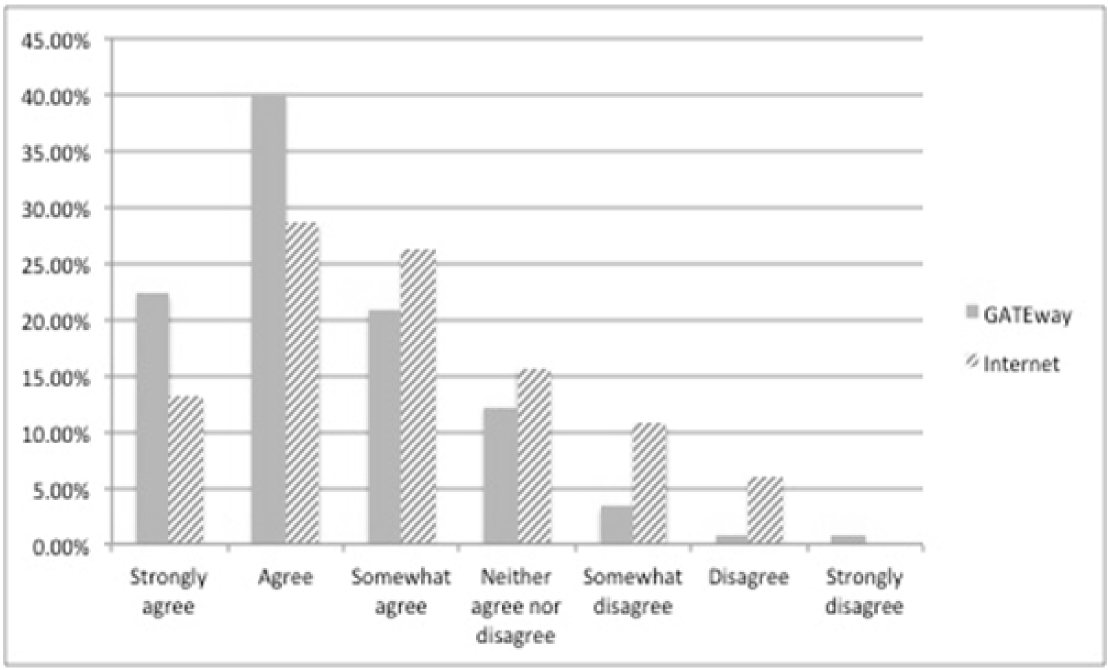

Figure 7.4. Responses to the statement “I can trust a driverless vehicle” separated by group. Reprinted from *Attitudes to autonomous vehicles,* by Hyde, Dalton and Stevens, 2017, Transport Research Laboratory (TRL). *ISBN: 978-1-910377-91-8*. Reprinted with permission.

The reduction in positivity in response to “I can trust a driverless vehicle” compared with “driverless cars are a good idea” perhaps indicates a general openness to the concept of the technology coupled with caution over the reliability of the systems as they currently exist. This is supported by the responses to the freeform response section at the end of the survey. Of the 57 participants who provided responses to this item, eight made the point that their level of trust in an autonomous vehicle would depend on the specifics of the vehicle, with factors such as manufacturer, safety record and independent reviews likely to influence their judgements. It therefore seems likely that people are open in principle to adopting this technology, but that their decisions when given the opportunity to use or purchase an autonomous vehicle will depend on the characteristics of the vehicle in question. This attitude is summed up by the following example comment:

> *“A lot of the trustworthiness of the driverless vehicles would depend on independent reviews, safety ratings, and other people's experiences, as well*

as the trustworthiness of the manufacturer and the technology.” - respondent comment

Main concerns about autonomous vehicles.

This survey did not ask directly about the nature of people’s concerns around autonomous vehicles. However, many respondents did set out their concerns in the freeform response section at the end of the survey. The most frequently raised concerns involved interactions between autonomous and non-autonomous vehicles in the early phases of introduction and the security of the software. Other common concerns involved: the ethical questions raised by the introduction of the technology; the likelihood of problems with the technology in the early phases of introduction; and the need for new policies and legislation to regulate the use of self-driving vehicles. Typical example comments were:

“Based on what we know about automated systems I would trust it more than a human being in a driving task. It's the other humans on the road that will cause problems in its initiation process.”- respondent comment.

“My big concern is the possibility of a hack attack.”- respondent comment.

“I am concerned about the inevitable situation where the vehicle chooses to kill its driver to save a greater number of external people in an unavoidable collision.”- respondent comment.

Different degrees of automation.

79% of people reported that they were extremely or moderately likely to want to be a passenger in a vehicle with semi-autonomous features (specified in the survey as: “e.g. lateral and longitudinal control, motorway assist systems for travel on high speed roads, remote control parking, Volvo or Tesla autopilot”), with only 3% being extremely or moderately unlikely. Fewer people (73%) were extremely or moderately likely to want to be a driver in a vehicle “which you can allow to take over driving”, with 10% reporting that they were extremely or moderately unlikely. The patterns of responding between these two questions were significantly different ($\chi2$ $(1, N = 464)$ = 8.65, p =.01), suggesting that there may currently be a higher degree of acceptance for partial automation than for full automation.

Effects of providing information.

Participants who received up-to-date information about autonomous vehicles were no more positive in their responses to the statements “Driverless cars are a good idea” and “I can trust a driverless vehicle” than participants who did not receive the information. This pattern was the same across the GATEway and internet groups (p >.60 for all comparisons). These findings are in line with those of Souders and Charness (2016) whose information manipulation had little effect on overall trust levels (although note that their study was in participants aged 55 and over, whereas the current study surveyed a broader range of ages).

Discussion

The attitudes towards autonomous vehicles revealed in this survey are broadly very positive. This may reflect an increasing openness among the general public to engage with driverless cars. However, it is important to note that 64% of the sample was drawn from a group of people who had registered to take part in trials relating to the GATEway project, perhaps indicating an unusually high level of interest in this topic area. Indeed, this group were significantly more positive than the internet group in their responses to the statements “Driverless cars are a good idea” and “I can trust a driverless vehicle”. Nevertheless, the responses remained highly positive in general, even with these participants removed.

Of course, the internet group were also self-selecting and might be considered as more engaged with technology than the public in general. So, a further caveat of the findings concerns the representativeness of the sample. For example, by comparison with another larger survey study (Schoettle & Sivak, 2014), the present sample had a higher proportion of male respondents (60% vs. 53%) and was overall somewhat younger and more likely to have a university degree. In addition, both of these surveys were carried out online, necessarily excluding individuals without internet access. Thus, caution should be applied in generalising the findings too far beyond the group reached here.

The fact that providing additional information concerning the technology did not significantly change people’s attitudes suggests that knowledge level may not be a central factor in determining people's views of autonomous vehicles, at least for the respondents reached by the current survey. However, very few of the participants had experienced real-world travel in an autonomous vehicle (there were only 37 positive responses to this question out of 233 respondents, meaning that a maximum of 16% of participants had experienced the technology). It is likely that this type of experience would have a much larger impact on attitudes than a brief information sheet. For this reason, it may be useful to re-run this survey both before and after people's first experience of a real-world driverless vehicle. The survey could also be used to track changing attitudes on a more general level, as autonomous vehicles become more widespread, and their benefits become more visible.

Finally, although this survey and the other research described in this chapter used a subjective measure of trust in automation, recent research has identified gaze behaviour as a possible means of measuring trust levels more objectively. In a simulator study, Hergeth, Lorenz, Vilimek, and Krem (2016) examined the gaze behaviour and self-reported trust in automation of 35 participants during periods of highly automated driving (SAE level 3/4) and during periods of completing a non-driving related task (a simple target matching task completed on an iPad). Trust in automation was probed with a simple question that asked, "On a scale from 0% to 100%, how much do you trust the system?" (Hergeth et al., 2016, pg 4) and was presented 8 times during the testing session. There was a consistent relationship between drivers' trust in automation and their gaze behaviour. Specifically, in the automated driving condition, higher levels of trust in the vehicle automation were associated with lower levels of monitoring the automation and the driving environment, as measured by the number of glances per second. Additionally, in the non-driving condition, a higher trust in automation was associated with lower frequency monitoring of the autonomous system. Interestingly as the trial session progressed the participants' trust in the autonomous system increased, such that 60% of participants self-reported an increased level of trust and the same participants 54% showed a decrease in their monitoring frequency. Thus overall, this study shows that in principle a driver's trust in automation can be inferred from their visual monitoring behaviour, and this could be an interesting avenue for future research on this topic.

Chapter 8 – General Discussion

Introduction

The experiments reported in this thesis have investigated several aspects of vehicle automation and their effects on visual cognition and situation awareness using a variety of different methods and populations. In Chapters 2-5, I used a laboratory-based task to compare the effects of dual task workload on hazard perception across three groups: younger experienced drivers; older drivers; and novice drivers. In Chapter 6 I investigated the specific effect of vehicle automation on visual attention and situation awareness using eye tracking and a realistic simulated driving environment. Finally, in Chapter 7 I examined people's trust and acceptance of AVs on UK roads using an online survey. In this chapter, I also used a manipulation to assess the effect of providing up to date information about AVs and their future benefits on people's perceptions of AVs. In the next section, I will briefly highlight the key findings from each Chapter, their strengths and limitations, and their implications for cognitive psychology, human factors research and vehicle design.

Key Findings, Evaluation and Implications

In Chapter 2 I examined the effect of task workload on the perception of driving hazards, using a dual task paradigm in which a simple (vs. choice) RT task was paired with a video-based hazard detection task. The results suggested that the young experienced drivers who took part in this study exhibited better hazard detection performance under dual (vs. single) task conditions. This result was surprising given the weight of evidence that shows a cost to primary task performance under dual (vs. single) task conditions (e.g. Ettwig & Bronkhorst, 2015; Pashler, Johnston, & Ruthruff, 2001; Pashler, 1994). The effect observed could be potentially explained by the 'attentional boost' effect as proposed by Swallow and Jiang (2010, 2011, 2012) which has demonstrated that detection of a target in one task can temporarily enhance processing of concurrently-presented (yet task-irrelevant) information. However, these findings would constitute a similar but longer-lasting effect. One implication of these results is that keeping drivers engaged in some element of driving (e.g. lane control) is likely to improve hazard response performance by comparison with full automation. This could be an important consideration for future human factors research and AV design.

The study design in this chapter was designed to control the visual cognitive demands of the task. By using a simplified primary driving task with a simple (vs. choice) response secondary task, I could look specifically at the effects of dual task cognitive workload, while all other aspects of the visual display remained constant. Furthermore, the use of official hazard perception videos meant that the stimuli were representative of real life driving hazards and scenarios (unlike the hazards presented in many simulator implementations). However, there are a few clear limitations to this study design. First, the use of the hazard perception videos and a non-driving related secondary task as a proxy for automation levels means that there is an issue with the generalisation of the results beyond this study. This proxy for automation lacks the realism of even a simulated automated driving task, in which multiple simultaneous tasks are required in order to control the vehicle. On balance this approach was taken due to the practical consideration of using a simpler, less realistic design to better control for any potential confounding variables. Next, the sample size used may have not been ideal and, compounded with the small effect size in the hazard detection RT, may indicate an overestimate of effect size and inflation in the type one error rate. The gold standard of 80% power was not achieved in this study and was observed to be 63% in the hazard detection RT analysis, which although not very low was under what would have been ideal. A power calculation was not conducted during the design of this study due to the lack of similar previous studies from which to get the required guide to expected effect size or alpha value. In the end, I based the sample size on those that had been used in previous studies into hazard perception and driving, which tended to use approximately 30 participants. Finally, related to the limitations raised above, the probability values observed for the hazard detection RT (p =.038) are only marginally significant, and do not suggest a very strong case to reject the null hypothesis. Therefore, the results of this chapter should be interpreted with caution, and, as will be discussed in the future research section, constitute a prime candidate for a replication study.

In Chapter 3 I examined the effect of task workload on the perception of driving hazards in older drivers, using the same paradigm as Chapter 2. The results revealed that older drivers' RTs for detecting driving hazards were on average 1100ms longer when they also completed a high load secondary task than when they completed the hazard detection task alone. Unlike the pattern of results seen in Chapter 2, this finding is in line with previous research on dual task interference, which typically find significant costs in dual (vs. single) task conditions. However, subsequent analysis indicated that there might have been a speed-accuracy trade-off, because the longer overall RTs to the hazards as task difficulty increased were accompanied by an increase in accuracy. In line with this possibility, there was no difference in inverse efficiency scores between the load conditions. This is suggestive of a strategy change by older drivers, such that they sacrifice speed as task load increases, to achieve better accuracy. One possible implication of these results is that the reduction in task demand associated with increasing levels of automation could benefit older people's driving performance and remove the need for older drivers to implement these types of trade-off. These are important considerations for AV designers and future human factors research, particularly in the areas of handover and takeover request research, as older drivers may have difficulties in quickly detecting and responding to a real driving hazard during an emergency takeover request, and further research is needed (as will be discussed below).

In Chapter 4 I used the same paradigm to examine the effect of task workload on the perception of driving hazards in novice drivers. The results demonstrated that RTs for detecting driving hazards were not significantly different as dual task workload increased. Additionally, there was no significant change in hazard detection accuracy also as workload increased. Interestingly, the improvement in hazard detection with the additional task that was seen in Chapter 2 with younger experienced drivers was not mirrored in the current findings, perhaps due to the relative inexperience of the novice drivers used in Chapter 4. Nevertheless, both experiments produced findings that contradict previous research on dual task interference, which typically finds significant costs in dual (vs.

single) task conditions. This indicates that novice drivers might be partially better protected from dual task interference potentially because of their younger age, and they might be able to switch between secondary task of varying workload with little drop in their primary task performance (and indeed, in the case of Chapter 2, some evidence for improvement). A reasonable explanation for these non-differences might stem from the participants having relatively high working memory capacity (WMC) due to the sample being taken from a university population. Higher WMC in novice drivers may confer a protective quality when completing dual or multiple tasks (e.g. Wood, Hartley, Furley, & Wilson, 2016; Feldhütter, Gold, Schneider, & Bengler, 2017). However, this idea was not examined directly in this study and would need further clarification in a follow up study.

In Chapter 5 I examined the differences between younger experienced, older and novice drivers in the effects of task demand on hazard detection performance. The pattern of results was mixed. The workload measures indicated that the perceived workload increased reliably between the load conditions and did not differ between the groups, demonstrating that the groups experienced a similar overall level of subjective workload. Yet, older drivers in the no load condition reported requiring significantly more effort than the young experienced driving age sample reported to complete the task, possibly reflecting an overall reduced capacity in the older drivers, such that they already experienced the hazard detection task alone as relatively demanding. Older drivers also made more mistakes than both the young experienced and novice age samples and took significantly longer to respond in the secondary number probe task. In terms of the hazard task, load condition and driving group had little effect on detection RTs. However, in the hazard perception accuracy, both the older and novice drivers showed worse hazard performance than the young experienced drivers, in line with the literature showing worse hazard perception in older (vs. young experienced) drivers and in novice (vs. young experienced) drivers (Caird et al., 2005; Crundall & Underwood, 1998; Horswill et al., 2015; Maltz & Shinar, 1999; Scialfa et al., 2011, 2012). This could suggest that for at risk driving groups, including both older drivers and novice, inexperienced drivers, additional hazard perception training might be beneficial.

In Chapter 6 I examined the effect of vehicle automation on non-emergency takeover performance and situation awareness. One of the most consistent findings was that takeover responses were quicker when the period of automated driving involved the occlusion of visual information from the driving scene by fog, compared with situations where this occlusion was not present. This finding was the same whether takeover response was measured in terms of the time taken to make a fixation back to the screen or the time taken to return the hands to the wheel. A practical implication of this result is that it might be feasible for vehicle designers and manufactures to deliberately accompany the takeover request with a stimulus that results in a high level of visual change (e.g. a large flashing box presented very briefly around the windscreen) in order to alert drivers in the way that the abrupt fog removal is thought to have done in my simulator study. The large individual differences in the takeover times observed might also lead to the suggestion that designers could present settings in the autonomous system to allow tailoring of the takeover request time to individual drivers' preferences. However, neither of these implications were examined directly in this work, and both would require future research before firm recommendations could be made.

The eye tracking analysis showed that in the first second after a takeover request visual attention was more dispersed in the with-fog compared to the no-fog condition. This demonstrates that when visual information has been reduced during automated driving, a driver is more likely to look holistically at the visual scene during the subsequent takeover process, rather than focussing on a central point where arguably most of the important task relevant information is available. This is indicative of a driver in the with-fog condition reorienting themselves back into the visual scene and the task of driving by widely scanning the environment. Importantly, this reorientation seems to be quick and resolves after 1 second. This may suggest that vehicle manufacturers could design cabins that allow the driver to not always be facing the roadway in front of them, knowing that potentially drivers can visually reorient themselves into the driving scene very effectively even when they have not been monitoring the scene during the automated period. However, this idea would need to be approached with utmost caution and would require further rigorous testing.

The driving behaviour measures showed that the braking RT, speed, headway, and TTC at the start of the driving hazard did not significantly differ between the two occlusion conditions. This indicates that in the approximately 3-4 seconds after the driver takes back control of the vehicle they appear to regain a similar amount of vehicular control whether or not they were monitoring the scene during the automated period. Again, a tentative implication for designers from these findings is that the ability to monitor the driving scene throughout the automated period might not be important in determining driving performance following a takeover request, at least under the conditions that were simulated in Chapter 6.

Responding to a hazard was observed to differ following automated driving as compared with manual driving in the current study. Specifically, participants tended to drive more slowly in the autonomous condition at the start of the driving hazard, perhaps suggesting that drivers are more cautious in the period following the handover and therefore adopt lower speeds. However, this cautious driving behaviour was not mirrored in the headway and TTC analysis. Additionally, braking in response to a developing hazard was no slower following an autonomous handover than following a period of manual driving.

Taken all together these results show that after a period of automation and a short 3-4 sec duration after takeover before the hazard starts, drivers in the autonomous conditions show very few differences in driving performance compared to manual drivers. This potentially indicates that drivers can quickly and effectively reengage with manual driving following a period of autonomous driving. However, as previously mentioned in Chapter 6, the driving environment used was benign and although drivers were able to maintain good vehicular control under these conditions, they might not have been ready for severe avoidance manoeuvres after a period of autonomous driving or in an emergency situation. The implication of this set of driving behaviour results is that there is little impact of automation and screen occlusion on situation awareness and this may cautiously inform on future vehicular design and human factors research.

The study reported in Chapter 6 has several strengths by comparison to the laboratory-based experiments presented in Chapters 2-5. First, the use of a driving simulator edges the methodology closer to a more ecologically valid methodology compared with the paradigm used in Chapters 2-5. This allowed me to make more generalisable claims and provide clear real-world interpretations of the results. Second, the task used is very like the takeover type request that drivers of level 2/3 automation are likely to experience in the future. Third, the use of a real-world additional task during the period of automation is a clear strength (and potential limitation). Electronic devices are likely to be the most frequent things with which drivers will interact during periods of automation, ahead of books, magazines, videos etc. Therefore, this gives the study a more realistic edge than if they were completing a controlled and contrived task (as seen in previous research e.g. Hergeth, Lorenz, Vilimek, & Krems, 2016). It is also a slight weakness, as the participants chose to use the device in whatever way they wanted. This means that the type of task (e.g. playing a game, watching videos, sending texts or browsing the internet) was not controlled for, and each task is likely to involve differing levels of cognitive or attentional involvement (e.g. having to reengage with driving whilst in the middle of writing a text/mail is potentially different from simply reading a news article or playing a computer game) which is likely to have added noise to the results.

There are also some limitations of the study. First, the simulator used was of a moderate fidelity, i.e. it was not as immersive as using a full vehicle based simulator or a real vehicle and task. However, this is only a slight limitation as the mini simulator set up utilised the excellent software, facilities and programming expertise at TRL to create a realistic and immersive driving task. Second, the recording fidelity of the eye tracker used was low. The eye tracker only recorded at a 20Hz data sampling rate, which is low when compared to more modern eye trackers that can record at a sample rate of 1000Hz (i.e. 1000 samples per second). The down side to this low sample rate in this study is reduced temporal accuracy.

Higher sampling rates produce better temporal accuracy when measuring the duration of fixations and saccades. Specifically, the average temporal accuracy error will be approximately half the duration of the time between samples. For example, a sampling rate of 1000 Hz will lead to an average error of 0.5 ms and a sampling rate of 60 Hz (i.e. samples every 16.7 ms) will lead to an average error of approximately 8 ms. However, given that I was not specifically looking at the number, or velocity of saccades, a low sample rate was deemed acceptable and suitable for collecting satisfactory fixation data.

In the final Chapter 7 I examined the trust and acceptance of AVs. The attitudes towards autonomous vehicles revealed in this survey were broadly very positive. This may reflect an increasing openness among the public to engage with driverless cars. However, it is important to note that 64% of the sample was drawn from a group of people who had registered to take part in trials relating to an automated vehicle project, perhaps indicating an unusually high level of interest in this topic area. Indeed, this group were significantly more positive than the internet group in their responses to the statements “driverless cars are a good idea” and “I can trust a driverless vehicle”. Nevertheless, the responses remained highly positive in general, even with these participants removed. The implications from this study are clear: 1) UK-based transport users are interested in seeing AVs on UK roads; and 2) there is a good level of trust in AVs. This last point is qualified by the finding that there was a more diverse range of responses concerning people’s level of trust in AVs and the sample surveyed already had a keen interest in AVs so may not be truly representative.

One of the main strengths of this survey is that it can easily be re-run in the future to assess trends in trust and acceptance of automation. However, the sample used in Chapter 7 had a higher proportion of male respondents (60% vs. 53%) and was overall somewhat younger and more likely to have a university degree than the general population. This may slightly weaken the generalisability of the results. Finally, there were only 233 useable respondents. In a future study, a far larger sample, in the thousands, would hopefully address some of the issues highlighted above and broaden the generalisability of the results to the wider population.

Future Research

The findings of the studies presented in this thesis do highlight some key areas for future research. The results in Chapter 2, in which hazard detection performance improved as secondary task workload increased, were clearly in contrast with a large amount of previous cognitive psychology literature. Therefore, it would be sensible to run a replication study using the same methodology and stimuli with a new sample of drivers. Doing so will either confirm and strengthen the results found or indicate that the original results may have occurred due to a type 1 error. Additionally, a new study using a secondary task which engages another modality (e.g. audition) could be an interesting area for future research. Auditory warning cues are already being used in many vehicles (e.g. to indicate lane departure, communicate a takeover request or announce a need for drivers to return their hands to the wheel in certain level 2 vehicles) so examination of the effects of auditory task load would be an interesting topic for future work. This would allow for the removal of extra visual information while maintaining the workload manipulation, and would also be useful in testing whether the patterns of results reported in Chapter 2 hold up when engaging a different modality or whether the effects are restricted to contexts in which both tasks involve visual information.

The results of Chapter 3 showed a more traditional dual task interference effect in the older drivers. Although this is not a hugely surprising result, it is potentially an important one, in that older drivers may benefit from the reduction of task workload on their driving performance. However, it is vitally important that the hand over and takeover performance of older drivers is examined in order to ascertain if the reduction in tasks impacts their performance if they are asked to resume control of semi-autonomous vehicle at short notice. This possibility can be examined easily, safely and with good validity with the use of a driving simulator, using a paradigm like that used in Chapter 6.

The results of Chapter 4 show a mixed pattern for the novice drivers and would therefore warrant replication in a follow-up study. Given that working memory capacity differences might have accounted for many aspects of the results (across Chapters 2-5, but in particular for novice drivers) controlling for and measuring WMC capacity of individual participants should be taken into account in future research.

In both Chapters 3 and 4, the literature presented shows that driver and perceptual training interventions for novice and older drivers can improve their hazard perception and visual scanning behaviour. It would be therefore an interesting future avenue for research to assess the impact of training interventions tailored to help all drivers with their takeover control and visual scanning behaviour after automation.

The results of Chapter 6 show an interesting pattern of results in takeover performance and situation awareness. A clear area for future research would be to take this study and apply it to the at-risk driving groups as previously described in Chapters 3 and 4. Doing so would help to tease apart the differences in driving performance in a more realistic setting and examine if the at-risk driving groups' takeover performance is different from a younger experienced population. This would allow human factors research to carefully inform policy decisions regarding the safety of level 2/3 automation across the whole population.

Finally, the results of the Chapter 7 showed an interesting pattern, with a clear area for future research: namely, repeating the survey on a periodic basis (e.g. biennial), with a larger sample (>1000 people) and open to all members of the public. This would provide a useful tool for assessing any trends in people's trust and acceptance of AVs over time.

Contribution to the Field

The contributions of this body of work to psychological and human factors knowledge is varied. First, for cognitive psychology, the surprising results of Chapter 2 show that under certain circumstances the addition of a secondary task can improve primary task performance. These findings may not relate directly to the attentional boost as it is currently defined, but may instead raise the possibility of a similar but longer-lasting effect. The results of Chapters 3-5 of the thesis show that there are some small key differences in task performance under dual task workload conditions in the at risk driving groups compared to younger experienced drivers, which could inform future human factors research. The results of Chapter 6 cautiously add to the knowledge of takeover performance, situation awareness and visual attention in younger experienced drivers after a period of vehicle automation. Finally, the results of Chapter 7 establish the public's levels of trust and acceptance of AVs in 2017, providing a potential avenue for assessing these perceptions as they develop into the future.

References

AA. (2017). There's no safe or unsafe age for a driver. Retrieved September 20, 2017, from https://www.theaa.com/driving-advice/safety/older-drivers

AAA, A. engineering. (2016). *Vehicle Technology survey*. Retrieved from https://publicaffairsresources.aaa.biz/wp-content/uploads/2016/02/Automotive-Engineering-ADAS-Survey-Fact-Sheet-FINAL-3.pdf

Adam, E. C. (1993). Fighter cockpits of the future. In *[1993 Proceedings] AIAA/IEEE Digital Avionics Systems Conference* (pp. 318–323). IEEE. https://doi.org/10.1109/DASC.1993.283529

Andersen, G. J., Cisneros, J., Saidpour, A., & Atchley, P. (2000). Age-related differences in collision detection during deceleration. *Psychology and Aging*, *15*(2), 241–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10879579

Andersen, G. J., & Enriquez, A. (2006). Aging and the detection of observer and moving object collisions. *Psychology and Aging*, *21*(1), 74–85. https://doi.org/10.1037/0882-7974.21.1.74

Annett, J. (2002). A note on the validity and reliability of ergonomics methods. *Theoretical Issues in Ergonomics Science*, *3*(2), 228–232. https://doi.org/10.1080/14639220210124067

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8

Bansal, P., & Kockelman, K. M. (2016). Are we ready to embrace connected and self-driving vehicles? A case study of Texans. *Transportation*, 1–35. https://doi.org/10.1007/s11116-016-9745-z

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *Vol. 57*(No 1), 289–300. Retrieved from http://www.jstor.org/stable/2346101?seq=1#page_scan_tab_contents

Bennett, P. J., Sekuler, R., & Sekuler, A. B. (2007). The effects of aging on motion detection and direction identification. *Vision Research*, *47*(6), 799–809. https://doi.org/10.1016/j.visres.2007.01.001

Borowsky, A., & Oron-Gilad, T. (2013). Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. *Accident; Analysis and Prevention*, *59*, 548–65. https://doi.org/10.1016/j.aap.2013.07.008

Borowsky, A., Shinar, D., & Oron-Gilad, T. (2010). Age, skill, and hazard perception in driving. *Accident; Analysis and Prevention*, *42*(4), 1240–9. https://doi.org/10.1016/j.aap.2010.02.001

Broadbent, D. E. (1952). Failures of attention in selective listening. *J Exp Psychol Hum Percept Perform*, *44*(6), 428–433.

Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review*, *64*(3), 205–215.

Bruyer, R., & Brysbaert, M. (2011). Combining Speed and Accuracy in Cognitive Psychology: Is the Inverse Efficiency Score (IES) a Better Dependent Variable than the Mean Reaction Time (RT) and the Percentage Of Errors (PE)? *Psychologica Belgica*, *51*(1), 5. https://doi.org/10.5334/pb-51-1-5

Butcher, L. (2016). *Young and novice drivers*. Retrieved from http://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN00517

Caird, J. K., Edwards, C. J., Creaser, J. I., & Horrey, W. J. (2005). Older driver failures of attention at intersections: using change blindness methods to assess turn decision accuracy. *Human Factors*, *47*(2), 235–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16170936

Cantin, V., Lavallière, M., Simoneau, M., & Teasdale, N. (2009). Mental workload when driving in a simulator: Effects of age and driving complexity. *Accident Analysis & Prevention*, *41*(4), 763–771. https://doi.org/10.1016/j.aap.2009.03.019

Cavoli, C., Phillips, B., Cohen, T., & Jones, P. (2017). Social and behavioural questions associated with automated vehicles: a literature review. Retrieved from https://trid.trb.org/view.aspx?id=1457834

CCAV. (2016). *Connected and Autonomous Vehicles: The future?* Retrieved from https://publications.parliament.uk/pa/ld201617/ldselect/ldsctech/115/115.pdf

Centers for Disease Control and Prevention. (2015). *Web-based Injury Statistics Query and Reporting System (WISQARS)*.

Centers for Disease Control and Prevention. (2017). Teen Drivers: Get the Fact. Retrieved April 25, 2017, from https://www.cdc.gov/motorvehiclesafety/teen_drivers/teendrivers_factsheet.html

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, *27*(8), 951–964. Retrieved from http://www.perceptionweb.com/abstract.cgi?id=p270951

Chapman, P., Underwood, G., & Roberts, K. (2002). Visual search patterns in trained and untrained novice drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*(2), 157–167. https://doi.org/10.1016/S1369-8478(02)00014-1

Choi, J. K., & Ji, Y. G. (2015). Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*, *31*(10), 692–702. https://doi.org/10.1080/10447318.2015.1070549

Clarke, D. D., Ward, P., & Truman, W. (2005). Voluntary risk taking and skill deficits in young driver accidents in the UK. *Accident; Analysis and Prevention*, *37*(3), 523–9. https://doi.org/10.1016/j.aap.2005.01.007

Clarke, D. D., Ward, P., Truman, W., & Bartle, C. (2009). *Collisions Involving Older Drivers: An In-depth Study*. London, England. Retrieved from http://webarchive.nationalarchives.gov.uk/20090417002224/http:/www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme5/rsrrno109.pdf

Colcombe, A. M., Kramer, A. F., Irwin, D. E., Peterson, M. S., Colcombe, S., & Hahn, S. (2003, June 1). Age-related effects of attentional and oculomotor capture by onsets and color singletons as a function of experience. *Acta Psychologica*. North-Holland. https://doi.org/10.1016/S0001-6918(03)00019-2

Crundall, D. E., & Underwood, G. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics*, *41*(4), 448–458. https://doi.org/10.1080/001401398186937

Crundall, D. E., & Underwood, G. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics*, *41*(4), 448–458. https://doi.org/10.1080/001401398186937

Crundall, D., Underwood, G., & Chapman, P. (1999). Driving experience and the functional field of view. *Perception*, *28*(9), 1075–1087. Retrieved from http://www.perceptionweb.com/abstract.cgi?id=p2894

Darling, K. A., & Helton, W. S. (2014). Dual-task interference between climbing and a simulated communication task. *Experimental Brain Research*, *232*(4), 1367–77. https://doi.org/10.1007/s00221-014-3855-7

Davis, J., Animashaun, A., Schoenherr, E., & McDowell, K. (2008). Evaluation of semi-autonomous convoy driving. *Journal of Field Robotics*, *25*(11–12), 880–897. https://doi.org/10.1002/rob.20263

de Waard, D. (1996). *The measurement of drivers' mental workload*. University of Groningen. Retrieved from https://www.rug.nl/research/portal/publications/the-measurement-of-drivers-mental-workload(0d430a43-552d-4f3a-9c3e-315ada87e165).html

de Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*. https://doi.org/10.1016/j.trf.2014.06.016

Department for Transport. (2015). *Reported Road Casualties Great Britain: 2014 Annual Report Moving Britain Ahead*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/467465/rrcgb-2014.pdf

Department for Transport. (2016). *Reported road casualties in Great Britain, main results: 2015*.

Department for Transport/Highways Agency. (2009). Traffi c Signs Manual Traffi c Safety Measures and Signs for Road Works and Temporary Situations. Retrieved from http://www.ukroads.org/webfiles/tsmchap8part1.pdf

Desmond, P. A., & Hoyes, T. W. (1996). Workload variation, intrinsic risk and utility in a simulated air traffic control task: Evidence for compensatory effects. *Safety Science*, *22*(1–3), 87–101. https://doi.org/10.1016/0925-7535(96)00008-2

Deutsch, J. A., & Deutsch, D. (1963). Some theoretical considerations. *Psychological Review*, *70*, 80–90.

Dewar, R., Kline, D., Scheiber, F., & Swanson, A. (1997). *SYMBOL SIGNING DESIGN FOR OLDER DRIVERS*. Retrieved from https://www.fhwa.dot.gov/publications/research/safety/94069/94069.pdf

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–68. https://doi.org/10.1146/annurev-psych-113011-143750

Dowiasch, S., Marx, S., Einhäuser, W., & Bremmer, F. (2015). Effects of aging on eye movements in the real world. *Frontiers in Human Neuroscience*, *9*(February), 1–12. https://doi.org/10.3389/fnhum.2015.00046

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology (London, England : 1953)*, *92 Part 1*, 53–

78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11802865

DVLA. (2015). *Full and provisional driving licences by age and gender*.

DVSA. (2014). Hazard Perception | The Official DVSA Learning Zone | DVSA. Retrieved September 20, 2017, from https://www.dvsalearningzone.co.uk/shop/hazard-perception

Endsley, M. R. (1988). Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society Annual Meeting*, *32*(2), 97–101. https://doi.org/10.1177/154193128803200221

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462–92. https://doi.org/10.1080/001401399185595

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology*, *79*(2, Pt.2), 1–16. https://doi.org/10.1037/h0026873

Eriksson, A., & Stanton, N. A. (2017). Takeover Time in Highly Automated Vehicles. *Human Factors*, 1872081668583. https://doi.org/10.1177/0018720816685832

Ernst and Young. (2013). *Autonomes Fahren – die Zukunft des Pkw-Marktes?* Retrieved from http://www.ey.com/Publication/vwLUAssets/Autonomes_Fahren_-_die_Zukunft_des_Pkw-Marktes/$FILE/EY-Autopilot-2013-Praesentation.pdf

Ettwig, J. F., & Bronkhorst, A. W. (2015). Attentional switches and dual-task interference. *PloS One*, *10*(3), e0118216. https://doi.org/10.1371/journal.pone.0118216

Feldhütter, A., Gold, C., Schneider, S., & Bengler, K. (2017a). *How the duration of automated driving influences take0over performance and gaze behavior*.

Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=G5AJDgAAQBAJ&oi=fnd&pg=PA309&ots=jwggwCHwdv&sig=P5MrN9JDbDzXCTwHvHEvkuNLEMI#v=onepage&q&f=false

Feldhütter, A., Gold, C., Schneider, S., & Bengler, K. (2017b). How the Duration of Automated Driving Influences Take - Over Performance and Gaze Behavior. In *Advances in ergonomic design of systems, products and processes* (pp. 309–318). Retrieved from https://www.researchgate.net/profile/Anna_Feldhuetter/publication/301287138_How_the_Duration_of_Automated_Driving_Influences_Take-Over_Performance_and_Gaze_Behavior/links/5710dc1608ae68dc790ae2e9.pdf

Galpin, A., Underwood, G., & Crundall, D. (2009). Change blindness in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(2), 179–185. https://doi.org/10.1016/j.trf.2008.11.002

Garcia, D., Kreutzer, C., Badillo-Urquiola, K., & Mouloua, M. (2015). Measuring Trust of Autonomous Vehicles: A Development and Validation Study (pp. 610–615). Springer, Cham. https://doi.org/10.1007/978-3-319-21383-5_102

GATEway. (2017). About | GATEway Project. Retrieved September 20, 2017, from https://gateway-project.org.uk/about/

Gkartzonikas, C., & Gkritza, K. (2017). *A Literature Review on Surveys for Autonomous Vehicles*. West Lafayette, IN. Retrieved from https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=TRF2017&paper_id=47

Glass, J. M., Schumacher, E. H., Lauber, E. J., Zurbriggen, E. L., Gmeindl, L., Kieras, D. E., & Meyer, D. E. (2000). Aging and the psychological refractory period: task-coordination strategies in young and old adults. *Psychology and Aging*, *15*(4), 571–95. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11144318

Godwin, H. J., Benson, V., & Drieghe, D. (2013). Using Interrupted Visual Displays to

Explore the Capacity, Time Course, and Format of Fixation Plans During Visual Search. *J Exp Psychol Hum Percept Perform*. https://doi.org/10.1037/a0032287

Godwin, H. J., Hyde, S., Taunton, D., Calver, J., Blake, J. I. R., & Liversedge, S. P. (2013). The Influence of Expertise on Maritime Driving Behaviour. *Applied Cognitive Psychology*, *27*(4), 483–492. https://doi.org/10.1002/acp.2925

Godwin, H. J., Liversedge, S. P., A, K. J., Boardman, M., Cornes, K., & Donnelly, N. (2015). The influence of experience upon information-sampling and decision-making behaviour during risk assessment in military personnel. *Visual Cognition*, *23*(4), 415–431. https://doi.org/10.1080/13506285.2015.1030488

Gold, C., Dambock, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1938–1942. https://doi.org/10.1177/1541931213571433

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, *52*(C), 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hartley, A. A., & Little, D. M. (1999). Age-related differences and similarities in dual-task interference. *Journal of Experimental Psychology. General*, *128*(4), 416–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10650582

Henderson, J. M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Percept Psychophys*, *45*, 196–208.

Hergeth, S., Lorenz, L., & Krems, J. F. (2017). Prior Familiarization With Takeover Requests Affects Drivers' Takeover Performance and Automation Trust. *Human Factors*, *59*(3), 457–470. https://doi.org/10.1177/0018720816678714

Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. F. (2016). Keep Your Scanners Peeled. *Human Factors*, *58*(3), 509–519. https://doi.org/10.1177/0018720815625744

Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. F. (2016). Keep Your Scanners Peeled:

Gaze Behavior as a Measure of Automation Trust During Highly Automated Driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(3), 509–519. https://doi.org/10.1177/0018720815625744

Hoc, J.-M., Young, M. S., & Blosseville, J.-M. (2009). Cooperation between drivers and automation: implications for safety. *Theoretical Issues in Ergonomics Science*, *10*(2), 135–160. https://doi.org/10.1080/14639220802368856

Hohenberger, C., Spörrle, M., & Welpe, I. M. (2016). How and why do men and women differ in their willingness to use automated cars? The influence of emotions across different age groups. *Transportation Research Part A: Policy and Practice*, *94*, 374–385. https://doi.org/10.1016/j.tra.2016.09.022

Horswill, M. S., Falconer, E. K., Pachana, N. A., Wetton, M., & Hill, A. (2015). The longer-term effects of a brief hazard perception training intervention in older drivers. *Psychology and Aging*, *30*(1), 62–7. https://doi.org/10.1037/a0038671

Horswill, M. S., Kemala, C. N., Wetton, M., Scialfa, C. T., & Pachana, N. A. (2010). Improving older drivers' hazard perception ability. *Psychology and Aging*, *25*(2), 464–9. https://doi.org/10.1037/a0017306

Horswill, M. S., Marrington, S. A., McCullough, C. M., Wood, J., Pachana, N. A., McWilliam, J., & Raikos, M. K. (2008). The hazard perception ability of older drivers. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *63*(4), P212–P218. Retrieved from http://psychsocgerontology.oxfordjournals.org/content/63/4/P212.abstract

Howard, D., & Dai, D. (2014). Public Perceptions of Self-Driving Cars: The Case of Berkeley, California. Retrieved from https://trid.trb.org/view.aspx?id=1289421

Hua Huang, C., Ming Luh, W., Sheu, C.-F., Tzeng, Y., & Chen, M. (2011). Analysis of Eye Movement Data Using Mixed Effects Modeling with a Poisson Regression Model. Retrieved from http://ir.lib.ncku.edu.tw/retrieve/124906/Analysis+of+eye+movements+data+Using+Mixed+Effects+Modeling+and+Poisson+Regression.pdf

Hyde, S., Dalton, P., & Stevens, A. (2017). *Attitudes to autonomous vehicles*.

London. Retrieved from https://trl.co.uk/reports/attitudes-autonomous-vehicles

Isler, R. B., Starkey, N. J., & Sheppard, P. (2011). Effects of higher-order driving skill training on young, inexperienced drivers' on-road driving performance. *Accident; Analysis and Prevention*, *43*(5), 1818–27. https://doi.org/10.1016/j.aap.2011.04.017

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2010). Foundations for an Empirically Determined Scale of Trust in Automated Systems. *http://dx.doi.org/10.1207/S15327566IJCE0401_04*.

John M, H. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504. https://doi.org/10.1016/j.tics.2003.09.006

Jonah, B. A., & Dawson, N. E. (1987). YOUTH AND RISK: AGE DIFFERENCES IN RISKY DRIVING, RISK PERCEPTION, AND RISK UTILITY. *ALCOHOL, DRUGS, AND DRIVING*, *3*(3–4). Retrieved from https://trid.trb.org/view.aspx?id=243005

Klein, C., Fischer, B., Hartnegg, K., Heiss, W. H., & Roth, M. (2000). Optomotor and neuropsychological performance in old age. *Experimental Brain Research*, *135*(2), 141–54. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11131498

Konstantopoulos, P., Chapman, P., & Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis & Prevention*, *42*(3), 827–834. https://doi.org/10.1016/j.aap.2009.09.022

KPMG. (2015). *Connected and Autonomous Vehicles – The UK Economic Opportunity*. Retrieved from https://www.smmt.co.uk/wp-content/uploads/sites/2/CRT036586F-Connected-and-Autonomous-Vehicles-–-The-UK-Economic-Opportu...1.pdf

Kuznetsova, A., Brockhoff, B., Haubo, R., & Maintainer, B. C. (2013). Package “lmerTest” Title Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://www2.uaem.mx/r-mirror/web/packages/lmerTest/lmerTest.pdf

Kyriakidis, M., Happee, R., & de Winter, J. C. F. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, *32*, 127–140. https://doi.org/10.1016/j.trf.2015.04.014

Lank, C., Haberstroh, M., & Wille, M. (2011). Interaction of Human, Machine, and Environment in Automated Driving Systems. *Transportation Research Record: Journal of the Transportation Research Board*, *2243*, 138–145. https://doi.org/10.3141/2243-16

Laubrock, J., Engbert, R., Rolfs, M., & Kliegl, R. (2007). Microsaccades Are an Index of Covert Attention. *Psychological Science*, *18*(4), 364–366. https://doi.org/10.1111/j.1467-9280.2007.01904.x

Lavallière, M., Simoneau, M., Tremblay, M., Laurendeau, D., & Teasdale, N. (2012). Active training and driving-specific feedback improve older drivers’ visual search prior to lane changes. *BMC Geriatrics*, *12*(1), 5. https://doi.org/10.1186/1471-2318-12-5

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N., & Dalton, P. (2014). Load Theory of Attention and Cognitive Control. In The Oxford Handbook of Attention. https://doi.org/10.1093/oxfordhb/9780199675111.013.003

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology. General*, *133*(3), 339–54. https://doi.org/10.1037/0096-3445.133.3.339

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception and Psychophysics*, *56*(2), 183–197. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7971119

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01

Locker, L., Hoffman, L., & Bovaird, J. A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods*, *39*(4), 723–730. https://doi.org/10.3758/BF03192962

Louw, T., Madigan, R., Carsten, O., & Merat, N. (2016). Were they in the loop during automated driving? Links between visual attention and crash potential. *Injury Prevention*, injuryprev-2016-042155. https://doi.org/10.1136/injuryprev-2016-042155

Louw, T., & Merat, N. (2017). Are you in the loop? Using gaze dispersion to understand driver visual attention during vehicle automation. *Transportation Research Part C: Emerging Technologies*, *76*, 35–50. https://doi.org/10.1016/j.trc.2017.01.001

Madden, D. J. (2007). Aging and Visual Attention. *Current Directions in Psychological Science*, *16*(2), 70–74. https://doi.org/10.1111/j.1467-8721.2007.00478.x

Madden, D. J., & Langley, L. K. (2003). Age-related changes in selective attention and perceptual load during visual search. *Psychology and Aging*, *18*(1), 54–67. https://doi.org/10.1037/0882-7974.18.1.54

Malsburg, von der, T. (2017). Saccades. Retrieved from https://github.com/tmalsburg/saccades

Maltz, M., & Shinar, D. (1999). Eye movements of younger and older drivers. *Human Factors*, *41*(1), 15–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10354803

Mäntylä, T., Karlsson, M. J., & Marklund, M. (2009). Executive Control Functions in Simulated Driving. *Applied Neuropsychology*, *16*(1), 11–18. https://doi.org/10.1080/09084280802644086

Maquestiaux, F., Didierjean, A., Ruthruff, E., Chauvel, G., & Hartley, A. (2013). Lost ability to automatize task performance in old age. *Psychonomic Bulletin &*

Review, *20*(6), 1206–12. https://doi.org/10.3758/s13423-013-0438-8

Maquestiaux, F., Hartley, A. A., & Bertsch, J. (2004). Can practice overcome age-related differences in the psychological refractory period effect? *Psychology and Aging*, *19*(4), 649–67. https://doi.org/10.1037/0882-7974.19.4.649

Maquestiaux, F., Laguë-Beauvais, M., Ruthruff, E., Hartley, A., & Bherer, L. (2010). Learning to bypass the central bottleneck: declining automaticity with advancing age. *Psychology and Aging*, *25*(1), 177–92. https://doi.org/10.1037/a0017122

Marottoli, R. A., Mendes de Leon, C. F., Glass, T. A., Williams, C. S., Cooney, L. M., Berkman, L. F., & Tinetti, M. E. (1997). Driving cessation and increased depressive symptoms: prospective evidence from the New Haven EPESE. Established Populations for Epidemiologic Studies of the Elderly. *Journal of the American Geriatrics Society*, *45*(2), 202–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9033520

Martens, M. H., Wilschut, E. S., & Pauwelussen MSc, J. (2008). Semi-Autonomous Driving: Do Drivers Still Respond to Unexpected Events? In *15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting*. Retrieved from http://trid.trb.org/view.aspx?id=906337

Maylor, E. A., & Lavie, N. (1998). The influence of perceptual load on age differences in selective attention. *Psychology and Aging*, *13*(4), 563–73. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9883457

McCarley, J. S., Steelman, K. S., & Horrey, W. J. (2014). The View from the Driver's Seat: What Good Is Salience? *Applied Cognitive Psychology*, *28*(1), 47–54. https://doi.org/10.1002/acp.2955

McDonald, C. C., Curry, A. E., Kandadai, V., Sommers, M. S., & Winston, F. K. (2014). Comparison of teen and adult driver crash scenarios in a nationally representative sample of serious crashes. *Accident Analysis & Prevention*, *72*, 302–308. https://doi.org/10.1016/j.aap.2014.07.016

McGwin, Jr, G., & Brown, D. B. (1999). Characteristics of traffic crashes among young, middle-aged, and older drivers. *Accident Analysis & Prevention*, *31*(3), 181–198. https://doi.org/10.1016/S0001-4575(98)00061-X

McKenna, F. P., & Crick, J. (1994). *Hazard perception in drivers: a methodology for testing and training*. Retrieved from http://www.trl.co.uk/reports-publications/report/?reportid=3121

McKenna, F. P., Horswill, M. S., & Alexander, J. L. (2006). Does anticipation training affect drivers' risk taking? *Journal of Experimental Psychology: Applied*, *12*(1), 1–10. https://doi.org/10.1037/1076-898X.12.1.1

Merat, N., & Jamson, A. H. (2009). Is drivers' situation awareness influenced by a fully automated driving scenario? In *Human Factors in Transport Security Academic Forum*. Retrieved from http://eprints.whiterose.ac.uk/84466/1/pMerat.pdf

Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., & Carsten, O. M. J. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 274–282. https://doi.org/10.1016/j.trf.2014.09.005

Merat, N., Jamson, a. H., Lai, F. C. H., & Carsten, O. (2012). Highly Automated Driving, Secondary Task Performance, and Driver State. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*(5). https://doi.org/10.1177/0018720812442087

Mok, B., Johns, M., Miller, D., & Ju, W. (2017). Tunneled In. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17* (pp. 2840–2844). New York, New York, USA: ACM Press. https://doi.org/10.1145/3025453.3025713

Monsell, Stephen and Driver, J. (2000). *Control of Cognitive Processes: Attention and Performance XVIII*. MIT Press. Retrieved from http://books.google.co.uk/books/about/Control_of_Cognitive_Processes.html?id=kO_baYlSVbwC&pgis=1

Morgan, P., & Alford, C. (2017). *VENTURER Trial 1: Planned Handover*. Bristol.

Retrieved from http://www.venturer-cars.com/wp-content/uploads/2017/05/VENTURER-Trial-1-Planned-Handover-Technical-Report.pdf

Mourant, R. R., & Rockwell, T. H. (1972). Strategies of Visual Search by Novice and Experienced Drivers. *Human Factors*, *14*(4), 325–335. https://doi.org/10.1177/001872087201400405

Murphy, G., & Greene, C. M. (2017). Load Theory Behind the Wheel; Perceptual & Cognitive Load Effects. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*. https://doi.org/10.1037/cep0000107

Naughton, K. (2017). Ford’s Dozing Engineers Side With Google in Full Autonomy Push - Bloomberg. Retrieved June 11, 2017, from https://www.bloomberg.com/news/articles/2017-02-17/ford-s-dozing-engineers-side-with-google-in-full-autonomy-push

NHTSA. (2013). U.S. Department of Transportation Releases Policy on Automated Vehicle Development. Retrieved from http://www.nhtsa.gov/About+NHTSA/Press+Releases/U.S.+Department+of+Transportation+Releases+Policy+on+Automated+Vehicle+Development

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. https://doi.org/10.1109/3468.844354

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244. https://doi.org/10.1037/0033-2909.116.2.220

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology*, *52*, 629–51. https://doi.org/10.1146/annurev.psych.52.1.629

Payre, W., Cestac, J., & Delhomme, P. (2016). Fully Automated Driving: Impact of

Trust and Practice on Manual Control Recovery. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(2), 229–241. https://doi.org/10.1177/0018720815612319

Petermeijer, S., Bazilinskyy, P., Bengler, K., & de Winter, J. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied Ergonomics*, *62*, 204–215. https://doi.org/10.1016/j.apergo.2017.02.023

Pollatsek, A., Narayanaan, V., Pradhan, A., & Fisher, D. L. (2006). Using eye movements to evaluate a PC-based risk awareness and perception training program on a driving simulator. *Human Factors*, *48*(3), 447–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17063961

Pollatsek, A., Romoser, M. R. E., & Fisher, D. L. (2012). Identifying and Remediating Failures of Selective Attention in Older Drivers. *Current Directions in Psychological Science*, *21*(1), 3–7. https://doi.org/10.1177/0963721411429459

Poulter, D. R., & Wann, J. P. (2013). Errors in motion processing amongst older drivers may increase accident risk. *Accident Analysis & Prevention*, *57*, 150–156. https://doi.org/10.1016/j.aap.2013.03.031

Quimby, A. R., & WATTS, G. R. (1981). HUMAN FACTORS AND DRIVING PERFORMANCE. *Publication of: Transport and Road Research Laboratory*. Retrieved from http://trid.trb.org/view.aspx?id=174634

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with “The Language-as-fixed-effect fallacy”: Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*(3), 416–426. https://doi.org/10.1006/jmla.1999.2650

RAC. (2013). Older driver numbers hit four million. Retrieved September 20, 2017, from http://www.racfoundation.org/media-centre/older-driver-numbers-hit-four-million

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How Traffic Situations and Non-Driving Related Tasks Affect the Take-Over Quality in

Highly Automated Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 2063–2067. https://doi.org/10.1177/1541931214581434

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19449261

Rayner, K., Castelhano, M. S., & Yang, J. (2009). Eye movements and the perceptual span in older and younger readers. *Psychology and Aging*, *24*(3), 755–760. https://doi.org/10.1037/a0014300

Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science (New York, N.Y.)*, *278*(5343), 1616–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9374459

Reid, G. B., & Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. *Advances in Psychology*, *52*(C), 185–218. https://doi.org/10.1016/S0166-4115(08)62387-0

Rogé, J., Ndiaye, D., & Vienne, F. (2014). Useful visual field training: A way to improve elderly car drivers' ability to detect vulnerable road users. *Transportation Research Part F: Traffic Psychology and Behaviour*, *26*, 246–257. https://doi.org/10.1016/j.trf.2014.08.005

RoSPA. (2010). *Older drivers policy paper*. London, England. Retrieved from http://www.rospa.com/rospaweb/docs/advice-services/road-safety/drivers/older-drivers-2010.pdf

Ross, L. A., Dodson, J. E., Edwards, J. D., Ackerman, M. L., & Ball, K. (2012). Self-rated driving and driving safety in older adults. *Accident; Analysis and Prevention*, *48*, 523–7. https://doi.org/10.1016/j.aap.2012.02.015

Ruiqi Ma. (2006). *The effects of in-vehicle automation and reliability on driver* *situation awareness and trust (Doctoral dissertation).* North Carolina State University, Raleigh, NC. Retrieved from http://repository.lib.ncsu.edu/ir/bitstream/1840.16/4155/1/etd.pdf

Ruiqi Ma, & Kaber, D. B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics*, *35*, 939–953. Retrieved from http://people.engr.ncsu.edu/dbkaber/papers/Ma_Kaber_IJIE_05.pdf

SAE. (2014). *Summary Of SAE International's Levels Of Driving Automation For On-ROAD Vehicles*. Retrieved from https://www.sae.org/misc/pdfs/automated_driving.pdf

Sagberg, F., & Bjørnskau, T. (2006). Hazard perception and driving experience among novice drivers. *Accident Analysis and Prevention*, *38*(2), 407–414. https://doi.org/10.1016/j.aap.2005.10.014

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying Fixations and Saccades in Eye-Tracking Protocols. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.68.2459&rep=rep1&type=pdf

Schall, M. C., Rusch, M. L., Lee, J. D., Dawson, J. D., Thomas, G., Aksan, N., & Rizzo, M. (2012). Augmented Reality Cues and Elderly Driver Hazard Perception. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*(3), 643–658. https://doi.org/10.1177/0018720812462029

Schoettle, B., & Sivak, M. (2014). A Survey of Public Opinion about Autonomous and Self-Driving Vehicles in the U . S ., the U . K ., and Australia, (UMTRI-2014-21), 2014. Retrieved from http://cdn.theatlantic.com/newsroom/img/posts/UMTRI-2014-21.pdf

Scialfa, C. T., Borkenhagen, D., Lyon, J., Deschênes, M., Horswill, M., & Wetton, M. (2012). The effects of driving experience on responses to a static hazard perception test. *Accident Analysis and Prevention*, *45*, 547–553. https://doi.org/10.1016/j.aap.2011.09.005

Scialfa, C. T., Deschênes, M. C., Ference, J., Boone, J., Horswill, M. S., & Wetton, M. (2011). A hazard perception test for novice drivers. *Accident Analysis and Prevention*, *43*(1), 204–208. https://doi.org/10.1016/j.aap.2010.08.010

Scott, H., Hall, L., Litchfield, D., & Westwood, D. (2013). Visual information search in simulated junction negotiation: Gaze transitions of young novice, young experienced and older experienced drivers. *Journal of Safety Research*, *45*, 111–116. https://doi.org/10.1016/j.jsr.2013.01.004

Seichepine, D. R., Neargarder, S., McCallum, M. E., Tabor, K., Riedel, T. M., Gilmore, G. C., & Cronin-Golomb, A. (2012). Luminance affects age-related deficits in object detection: implications for computerized psychological assessments. *Psychology and Aging*, *27*(2), 522–8. https://doi.org/10.1037/a0025576

Sekuler, A. B., Bennett, P. J., & Mamelak, M. (2000). Effects of aging on the useful field of view. *Experimental Aging Research*, *26*(2), 103–20. https://doi.org/10.1080/036107300243588

Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20. https://doi.org/10.1016/j.tics.2004.11.006

Sivak, M. (1996). The information that drivers use: is it indeed 90% visual? *Perception*, *25*(9), 1081–1089. https://doi.org/10.1068/p251081

Souders, D., & Charness, N. (2016). Challenges of Older Drivers' Adoption of Advanced Driver Assistance Systems and Autonomous Vehicles (pp. 428–440). https://doi.org/10.1007/978-3-319-39949-2_41

Strand, N., Nilsson, J., Karlsson, I. C. M., & Nilsson, L. (2014). Semi-automated versus highly automated driving in critical situations caused by automation failures. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 218–228. https://doi.org/10.1016/j.trf.2014.04.005

Swallow, K., & Jiang, Y. (2011). The role of timing in the attentional boost effect. *Attention, Perception, & Psychophysics*, *73*(2), 389–404. https://doi.org/10.3758/s13414-010-0045-y

Swallow, K. M., & Jiang, Y. V. (2010). The Attentional Boost Effect: Transient increases in attention to one task enhance performance in a second task. *Cognition*, *115*(1), 118–32. https://doi.org/10.1016/j.cognition.2009.12.003

Swallow, K. M., & Jiang, Y. V. (2011). The role of timing in the attentional boost effect. *Attention, Perception & Psychophysics*, *73*(2), 389–404. https://doi.org/10.3758/s13414-010-0045-y

Swallow, K. M., & Jiang, Y. V. (2012). Goal-relevant events need not be rare to boost memory for concurrent images. *Attention, Perception & Psychophysics*, *74*(1), 70–82. https://doi.org/10.3758/s13414-011-0227-2

Swallow, K. M., & Jiang, Y. V. (2014). Perceptual load and attentional boost: a study of their interaction. *Journal of Experimental Psychology. Human Perception and Performance*, *40*(3), 1034–45. https://doi.org/10.1037/a0035312

Tennant, C., Howard, S., Franks, B., Bauer, M. W., Pansegrau, P., Styśko-Kunkowska, M., … Cuevas-Badallo, A. (2015). *Autonomous Vehicles - Negotiating a Place on the Road*. Retrieved from http://www.lse.ac.uk/website-archive/newsAndMedia/PDF/AVs-negociating-a-place-on-the-road-1110.pdf

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786. https://doi.org/10.1037/0033-295x.113.4.766

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*(3), 282–299.

Treisman, A. M. (1964). Selective Attention In Man. *British Medical Bulletin*, *20*(1), 12–16. https://doi.org/10.1093/oxfordjournals.bmb.a070274

Trick, G. L., & Silverman, S. E. (1991). Visual sensitivity to motion: age-related changes and deficits in senile dementia of the Alzheimer type. *Neurology*, *41*(9), 1437–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1891094

Tsang, P. S., & Velazquez, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, *39*(3), 358–381. https://doi.org/10.1080/00140139608964470

Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. *Ergonomics*, *50*(8), 1235–1249. https://doi.org/10.1080/00140130701318707

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: sequences of eye fixations made by experienced and novice drivers. *Ergonomics*, *46*(6), 629–46. https://doi.org/10.1080/0014013031000090116

Underwood, G., Phelps, N., Wright, C., van Loon, E., & Galpin, A. (2005). Eye fixation scanpaths of younger and older drivers in a hazard perception task. *Ophthalmic & Physiological Optics : The Journal of the British College of Ophthalmic Opticians (Optometrists)*, *25*(4), 346–56. https://doi.org/10.1111/j.1475-1313.2005.00290.x

Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: a review of meta-analyses. *Neuroscience and Biobehavioral Reviews*, *26*(7), 849–57. https://doi.org/10.1016/S0149-7634(02)00071-4

Verster, J. C., & Roth, T. (2011). Standard operation procedures for conducting the on-the-road driving test, and measurement of the standard deviation of lateral position (SDLP). *International Journal of General Medicine*, *4*, 359–71. https://doi.org/10.2147/IJGM.S19639

Victor, T. W., Harbluk, J. L., & Engström, J. A. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(2), 167–190. https://doi.org/10.1016/j.trf.2005.04.014

Warren, W. H., Blackwell, A. W., & Morris, M. W. (1989). Age differences in perceiving the direction of self-motion from optical flow. *Journal of Gerontology*, *44*(5), P147-53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2768773

Watanabe, K., & Funahashi, S. (2014). Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. *Nature Neuroscience*, *17*(4), 601–11. https://doi.org/10.1038/nn.3667

Waymo. (2017). Waymo. Retrieved September 20, 2017, from https://waymo.com

Wells, P., Tong, S., Sexton, B., Grayson, G., & Jones, E. (2008). *Cohort II: A study of learner and new drivers.* London, United Kingdom: Department for Transport.

Winter, J. C. F. De, Stanton, N. A., Price, J. S., & Mistry, H. (2016). The effects of driving with different levels of unreliable automation on self-reported workload and secondary task performance. *International Journal of Vehicle Design*, *70*(4), 297. https://doi.org/10.1504/IJVD.2016.076736

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology-Human Perception and Performance*, *29*(2), 483–502.

Wood, G., Hartley, G., Furley, P. A., & Wilson, M. R. (2016). Working Memory Capacity, Visual Attention and Hazard Perception in Driving. *Journal of Applied Research in Memory and Cognition*. https://doi.org/10.1016/j.jarmac.2016.04.009

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics*, *58*(1), 1–17. https://doi.org/10.1080/00140139.2014.956151

Young, M. S., & Stanton, N. A. (2002a). Attention and automation: New perspectives on mental underload and performance. *Theoretical Issues in Ergonomics Science*, *3*(2), 178–194.

https://doi.org/10.1080/14639220210123789

Young, M. S., & Stanton, N. A. (2002b). Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. *Human Factors*, *44*(3), 365–75. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12502155

Young, M. S., & Stanton, N. A. (2007). Back to the future: brake reaction times for manual and automated vehicles. *Ergonomics*, *50*(1), 46–58. https://doi.org/10.1080/00140130600980789

Yurovsky, D., & Frank, M. C. (2017). Beyond naïve cue combination: salience and social cues in early word learning. *Developmental Science*, *20*(2), e12349. https://doi.org/10.1111/desc.12349

Appendix A

Mean distance in centimetres from the lane centre and standard deviation as a function of time bin (0-60 sec) and fog condition.

	No fog		With fog		
Time bins (sec)	Mean	SD	Mean	SD	Significance
0-1	1.2	1.7	1.2	1.6	.97 ^{ns}
1-2	1.7	2.0	1.8	2.0	.99 ^{ns}
2-3	0.9	1.0	0.9	1.0	.96 ^{ns}
3-4	3.6	5.6	3.5	5.8	<.0001***
4-5	27.6	25.1	23.9	25.0	<.0001***
5-6	64.3	43.9	53.4	38.0	<.0001***
6-7	69.3	43.4	62.1	36.7	<.0001***
7-8	63.3	41.2	64.5	40.0	.36 ^{ns}
8-9	61.4	43.5	64.4	42.6	.03*
9-10	62.1	44.4	59.1	38.9	.03*
10-11	56.3	39.3	55.1	36.6	.36 ^{ns}
11-12	53.5	38.7	51.6	34.3	.17 ^{ns}
12-13	51.5	34.7	48.7	35.8	.04*
13-14	51.9	34.9	46.5	36.6	<.0001***
14-15	51.8	31.6	48.6	34.8	.02*
15-16	50.0	30.4	48.5	35.6	.29 ^{ns}
16-17	48.4	30.0	46.1	33.0	.11 ^{ns}
17-18	49.1	34.2	45.9	31.9	.02 *
18-19	48.4	33.8	48.6	31.6	0.90 ^{ns}
19-20	48.9	34.2	48.0	31.8	.55 ^{ns}
20-21	48.6	32.9	47.4	33.4	.41 ^{ns}
21-22	48.3	32.1	49.8	33.4	.25 ^{ns}

	No fog		With fog		
Time bins (sec)	Mean	SD	Mean	SD	Significance
22--23	52.8	34.4	56.0	39.1	.02*
23-24	58.6	42.0	60.3	41.4	.20 ns
24-25	64.6	44.0	62.2	42.7	.08 ns
25-26	72.1	41.7	62.7	40.9	<.0001***
26-27	72.0	43.2	64.2	40.4	<.0001***
27-28	71.0	43.6	67.9	40.9	.02*
28-29	67.5	42.0	69.1	40.3	0.24 ns
29-30	66.2	42.0	68.4	39.3	.12 ns
30-31	63.9	40.2	68.5	38.8	<.0001***
31-32	62.8	39.4	70.3	38.7	<.0001***
32-33	64.7	39.1	68.6	38.5	.004**
33-34	71.7	43.1	65.0	37.9	<.0001***
34-35	70.3	39.4	63.5	38.6	<.0001***
35-36	64.7	38.3	61.6	36.0	.02*
36-37	60.7	38.0	60.6	35.6	0.95 ns
37-38	59.7	38.0	61.7	35.9	.14 ns
38-39	61.0	38.4	61.0	35.4	.97 ns
39-40	61.5	37.0	59.6	36.0	.19 ns
40-41	60.8	35.9	58.6	35.7	.11 ns
41-42	61.1	33.4	57.5	32.3	.009**
42-43	62.8	36.8	57.9	33.7	.0004***
43-44	63.3	39.0	59.5	36.1	.005**
44-45	62.9	39.8	59.1	35.8	.006**
45-46	62.6	38.5	56.1	33.5	<.0001***
46-47	61.3	37.8	52.4	32.6	<.0001***

	No fog		With fog		
Time bins (sec)	Mean	SD	Mean	SD	Significance
47-48	59.8	35.5	50.2	32.0	<.0001***
48-49	56.5	33.5	51.4	34.1	<.0001***
49-50	53.9	33.3	51.5	34.9	.08 ^{ns}
50-51	53.4	34.3	51.7	35.4	0.24 ^{ns}
51-52	52.9	34.0	52.0	36.0	.51 ^{ns}
52-53	53.2	33.2	54.7	37.6	.27 ^{ns}
53-54	52.7	32.6	55.7	35.4	.03*
54-55	51.7	30.8	55.8	34.6	.002**
55-56	50.3	30.4	54.1	33.7	.004**
56-57	49.3	30.1	52.8	34.0	.01*
57-58	49.8	31.2	51.8	32.5	.15 ^{ns}
58-59	52.0	33.1	53.8	33.1	.18 ^{ns}
59-60	53.6	33.5	57.0	35.9	.009**
60	53.3	32.4	57.8	36.5	.31 ^{ns}

(^{ns}= non-significant, *p<.05, **p<.01, ***p<.0001)

Appendix B

All questionnaire items and responses by group

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
Age (in years)	18 to 29	30 (20.3%)	29 (34.1%)	59 (25.3%)
	30 to 39	45 (30.4%)	29 (34.1%)	74 (31.7%)
	40 to 49	45 (30.4%)	11 (12.9%)	56 (24.0%)
	50 to 59	14 (9.5%)	8 (9.4%)	22 (9.4%)
	60 to 69	10 (6.75%)	4 (4.7%)	14 (6.0%)
	70 or older	4 (2.7%)	4 (4.7%)	8 (3.4%)
Gender	Male	117 (79%)	50 (60.2%)	167 (71.7%)
	Female	30 (20.3%)	32 (38.6%)	62 (26.6%)
	Trans gender or intersex	1 (0.7%)	0 (0%)	1 (0.4%)
	Would rather not disclose	0 (0%)	1(1.2%)	1 (0.4%)
Region of birth	UK	117 (79.1%)	57 (68.7%)	174 (75.3%)
	ROI	0 (0%)	3 (3.6%)	3 (1.3%)
	EU	7 (4.7%)	7 (8.4%)	14 (6.1%)
	Non-EU	24 (16.2%)	16 (19.3%)	40 (17.3%)
Region of driving licence	UK	114 (77.0%)	57 (67.9%)	171 (74.0%)
	ROI	1 (0.7%)	1 (1.2%)	2 (0.8%)
	EU	5 (3.4%)	3 (3.6%)	8 (3.4%)
	Non-EU	6 (4.1%)	3 (3.6%)	9 (3.9%)
	I don't drive	22 (14.9%)	20 (23.8%)	44 (18.9%)
Licence type (including transmission type permitted to drive)	Provisional manual	59 (30.9%)	29 (31.2%)	88 (30.9%)
	Provisional automatic	7 (3.7%)	3 (3.2%)	10 (3.5%)
	Full manual	118 (61.8%)	59 (63.4%)	177 (62.3%)

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
	Full automatic	7 (3.7%)	2 (2.1%)	9 (3.2%)
Length of time driving licence held (in years)	0-5 years	24 (19.2%)	12 (19.7%)	36 (19.3%)
	6-10 years	20 (16.0%)	15 (24.6%)	35 (18.8%)
	11-15 years	14 (11.2%)	10 (16.4%)	24 (12.9%)
	16-20 years	14 (11.2%)	6 (9.8%)	20 (10.7%)
	21-25 years	15 (12.0%)	5 (8.2%)	20 (10.7%)
	26-29 years	12 (9.6%)	1 (1.6%)	13 (6.9%)
	30+ years	26 (20.8%)	12 (19.7%)	38 (20.4%)
Average mileage driven per year (approximately)	Less than 1000	39 (31.2%)	14 (22.6%)	53 (29.8%)
	1001 to 2500	15 (12.0%)	10 (16. 9%)	15 (8.4%)
	2501 to 5000	18 (14.4%)	7 (11.3%)	25 (14.0%)
	5001 to 7500	12 (9.6%)	6 (9.7%)	18 (10.1%)
	7501 to 10000	12 (9.6%)	12 (19.4%)	24 (13.5%)
	10001 to 12500	14 (11.2%)	7 (11.3%)	21 (11.8%)
	12501 to 15000	7 (5.6%)	5 (8.1%)	13 (7.3%)
	15001 to 17500	4 (3.2%)	0 (0.0%)	4 (2.2%)
	17501 to 20000	2 (1.6%)	0 (0.0%)	2 (1.1%)
	20001 to 25000	0 (0.0%)	0 (0.0%)	0 (0.0%)
	25001 to 30000	1 (0.8%)	1 (1.6%)	2 (1.1%)
	30001 to 35000	1 (0.8%)	0 (0.0%)	1 (0.7%)
	35001 to 40000	0 (0.0%)	0 (0.0%)	0 (0.0%)
	40000 or more	0 (0.0%)	0 (0.0%)	0 (0.0%)
Have you experienced any driving aids in your time driving? (Select all applicable to you)	Adaptive cruise control	19 (22.1%)	33 (32.7%)	52 (28.8%)
	Lane-keep assistance	9 (10.4%)	15 (14.8%)	24 (12.8%)
	Parking assist	32 (37.2%)	31 (30.7%)	63 (33.7%)
	Collision			

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
	avoidance systems	12 (13.9%)	10 (9.9%)	22 (11.8%)
	Blind-spot monitoring	8 (9.3%)	9 (8.9%)	17 (.1%)
	Tesla autopilot	3 (3.5%)	0 (0%)	3 (1.6%)
	Other system not mentioned above	3 (3.5%)	3 (2.9%)	6 (3.2%)
Have you experienced travel in a self-driving vehicle? (Select all applicable to you)	Greenwich shuttle	1 (4.2%)	1 (16.7%)	2 (6.7%)
	Heathrow ultra-PODs	17 (71.8%)	3 (50%)	20 (66.7%)
	Google car	0 (0%)	0 (0%)	0 (0%)
	Other self-driving vehicle	6 (25%)	2 (33.3%)	8 (26.6%)
Overall do you enjoy driving?	Definitely yes	58 (46.0%)	30 (47.6%)	88 (46.6%)
	Probably yes	31 (24.6%)	18 (28.6%)	49 (25.9%)
	Neither yes or no	12 (9.5%)	9 (14.3%)	21 (11.1%)
	Probably not	13 (10.3%)	2 (3.2%)	15 (7.9%)
	Definitely not	12 (9.5%)	4 (6.3%)	16 (8.5%)
What roads do you typically drive on most of the time? (Select all applicable to you)	Urban	65 (32.9%)	66 (28.2%)	111 (25.7%)
	Suburban	50 (25.4%)	35 (14.9%)	72 (16.7%)
	Motorway	45 (22.8%)	61 (26.1%)	111 (25.7%)
	Rural	37 (18.8%)	72 (30.8%)	137 (58.5%)
What time(s) of day do you usually drive? (Select all applicable to you)	Early morning	30 (15.3%)	30 (12.9%)	60 (14.1%)
	Morning rush hour	28 (14.3%)	36 (15.5%)	64 (14.9%)
	Off-peak day time	63 (32.3%)	69 (29.7%)	132 (30.9%)
	Evening rush hour	37 (18.9%)	46 (19.8%)	83 (19.4%)
	Night time	37 (18.9%)	51 (21.9%)	88 (20.6%)

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
When traveling by motor vehicle are you typically the driver or passenger?	Driver	87 (58.8%)	45 (55.6%)	132 (57.6%)
	Passenger	61 (41.2%)	36 (44.4%)	97 (42.4%)
Are you registered disabled?	Yes	5 (3.4%)	4 (4.8%)	9 (3.9%)
	No	142 (6.6%)	77 (92.8%)	219 (95.2%)
	Would rather not say	0 (0.0%)	2 (2.4%)	2 (0.9%)
What is your occupation?	Professional	48 (32.4%)	29 (35.4%)	97 (40.4%)
	Clerical/office	12 (8.1%)	6 (7.3%)	18 (7.5%)
	Service worker	3 (2%)	1 (1.2%)	4 (1.7%)
	Executive/director	23 (15.5%)	15 (18.3%)	38 (15.8%)
	Sales worker	4 (2.7%)	0 (0%)	4 (1.7%)
	Skilled trade	3 (2%)	0 (0%)	3 (1.2%)
	Unskilled/laborer	0 (0%)	1 (1.2%)	1 (0.4%)
	Semi-skilled	2 (1.4%)	0 (0%)	2 (0.8%)
	IT professional	22 (14.9%)	10 (12.2%)	32 (13.3%)
	Student	5 (3.4%)	13 (15.9%)	18 (7.5%)
	Business owner	4 (2.7%)	0 (0%)	4 (1.7%)
	Retired	14 (9.5%)	3 (3.7%)	7 (2.9%)
	Other	8 (5.4%)	4 (4.9%)	12 (5.0%)
What is your highest level of education?	University or college degree	103 (70.1%)	57 (68.7%)	160 (69.6%)
	Other University or college qualification	19 (12.9%)	6 (7.2%)	25 (10.9%)
	Upper secondary school	12 (8.2%)	15 (18.1%)	27 (11.7%)
	Lower secondary			

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
	school qualification	13 (8.8%)	3 (3.6%)	16 (6.9%)
	None of these	0 (0%)	2 (2.4%)	2 (0.9%)
Where did you hear about this survey?	GATEway project			148 (63.5%)
	Social media			14 (6.0%)
	Email			43 (18.4%)
	Through RHUL			11 (4.7%)
	Through TRL			0 (0.0%)
	Other			17 (7.3%)
How likely are you to want be a driver in a vehicle which you can allow to take over driving	Extremely likely	81 (54.7%)	33 (38.8%)	114 (48.9%)
	Moderately likely	32 (21.6%)	23 (27.1%)	55 (23.6%)
	Slightly likely	10 (6.8%)	14 (16.5%)	24 (10.3%)
	Neither likely nor unlikely	6 (4.1%)	4 (4.7%)	10 (4.3%)
	Slightly unlikely	3 (2%)	3 (3.5%)	6 (2.6%)
	Moderately unlikely	5 (3.4%)	4 (4.7%)	9 (3.9%)
	Extremely unlikely	11 (7.4%)	4 (4.7%)	15 (6.4%)
How likely are you to want be a passenger in a vehicle with semi-autonomous features?	Extremely likely	95 (64.2%)	37 (44.6%)	132 (59.5%)
	Moderately likely	30 (20.3%)	22 (26.5%)	55 (24.7%)
	Slightly likely	10 (6.8%)	14 (16.9%)	24 (10.8%)
	Neither likely nor unlikely	7 (4.7%)	5 (6%)	13 (5.8%)
	Slightly unlikely	0 (0%)	3 (3.6%)	3 (1.3%)
	Moderately unlikely	1 (0.7%)	2 (2.4%)	3 (1.3%)
	Extremely unlikely	5 (3.4%)	0 (0%)	5 (2.2%)

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
In the future when driverless vehicle's are available as an option how likely are you to want to ride in one(regardless of cost)?	Extremely likely	108 (73.0%)	42 (50.6%)	150 (64.9%)
	Moderately likely	30 (20.3%)	22 (26.5%)	52 (22.5%)
	Slightly likely	5 (3.4%)	8 (9.6%)	13 (5.6%)
	Neither likely nor unlikely	4 (2.7%)	3 (3.6%)	7 (3.0%)
	Slightly unlikely	0 (0.0%)	3 (3.6%)	3 (1.3%)
	Moderately unlikely	0 (0.0%)	5 (6.0%)	5 (2.2%)
	Extremely unlikely	1 (0.7%)	0 (0.0%)	1 (0.4%)
Do you think that a human being should always be in charge of a vehicle?	Strongly agree	6 (4.1%)	5 (6.0%)	11 (4.8%)
	Agree	12 (8.1%)	12 (14.5%)	24 (10.4%)
	Somewhat agree	28 (18.9%)	20 (24.1%)	48 (20.8%)
	Neither agree nor disagree	18 (12.2%)	13 (15.7%)	31 (13.4%)
	Somewhat disagree	11 (7.4%)	10 (12.0%)	21 (9.1%)
	Disagree	30 (20.3%)	17 (20.5%)	47 (20.3%)
	Strongly disagree	43 (29.1%)	6 (7.2%)	49 (21.2%)
How much do you agree with the following statement: 'driverless cars are a good idea?'	Strongly agree	98 (66.2%)	37 (44.6%)	135 (58.19%)
	Agree	29(19.6%)	24 (28.9%)	53 (22.84%)
	Somewhat agree	17(11.5%)	14 (16.9%)	32 (13.79%)
	Neither agree nor disagree	1 (0.7%)	4 (4.8%)	5 (2.16%)
	Somewhat disagree	2 (1.4%)	3 (3.6%)	5 (2.16%)
	Disagree	0 (0%)	1 (1.2%)	1 (0.43%)
	Strongly disagree	1 (0.7%)	0 (0%)	1 (0.43%)
How much do you agree with the following	Strongly agree	117 (79.1%)	53 (63.9%)	170 (73.6%)
	Agree	18 (12.2%)	12 (14.5%)	30 (13.0%)

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
statement: 'driverless cars are an exciting prospect?	Somewhat agree	9 (6.1%)	10 (12.0%)	19 (8.2%)
	Neither agree nor disagree	1 (0.7%)	4 (4.8%)	5 (2.2%)
	Somewhat disagree	2 (1.4%)	3 (3.6%)	5 (2.2%)
	Disagree	1 (0.7%)	1 (1.2%)	2 (0.9%)
	Strongly disagree	0 (0.0%)	0 (0.0%)	0 (0.0%)
Do you think driverless cars should take control to prevent a crash?	Definitely yes	88 (59.5%)	34 (41.0%)	122 (52.8%)
	Probably yes	47 (31.8%)	37 (44.6%)	84 (36.4%)
	Might or might not	7 (4.7%)	11 (13.3%)	18 (7.8%)
	Probably not	4 (2.7%)	0 (0.0%)	4 (1.7%)
	Definitely not	2 (1.4%)	1 (1.2%)	3 (1.3%)
How much do you agree with the following statement: 'If 90% or more of accidents are down to human error then there is a strong case for taking driver control out of the equation	Strongly agree	89 (60.1%)	30 (36.1%)	119 (51.5%)
	Agree	29 (19.6%)	23 (27.7%)	52 (22.5%)
	Somewhat agree	19 (12.8%)	23 (27.7%)	42 (18.2%)
	Neither agree nor disagree	2 (1.4%)	2 (2.4%)	4 (1.7%)
	Somewhat disagree	4 (2.7%)	3 (3.6%)	7(3.0%)
	Disagree	5 (3.4%)	2 (2.4%)	7(3.0%)
	Strongly disagree	0 (0.0%)	0 (0.0%)	0 (0.0%)
How much do you agree with the following statement: 'advances in engineering sciences and automotive technology will	Strongly agree	89 (60.1%)	33 (39.8%)	122 (52.8%)
	Agree	35 (23.6%)	27 (32.5%)	62 (26.8%)
	Somewhat agree	16 (10.8%)	13 (15.7%)	29 (12.6%)
	Neither agree nor disagree	5 (3.4%)	6 (7.2%)	11 (4.8%)
	Somewhat	1 (0.7%)	4 (4.8%)	5 (2.2%)

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
allow driverless cars to be at least as safe as human drivers'	disagree	2 (1.4%)	0 (0.0%)	2 (0.9%)
	Disagree	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Strongly disagree			
How much do you agree with the following statement: 'driverless cars may be suitable for use in other countries (e.g. USA) but they're not suitable for use on our roads'	Strongly agree	1 (0.7%)	0 (0.0%)	1 (0.4%)
	Agree	3 (2.0%)	5 (6.0%)	8 (3.5%)
	Somewhat agree	8 (5.4%)	5 (6.0%)	13 (5.6%)
	Neither agree nor disagree	9 (6.1%)	11 (13.3%)	20 (8.7%)
	Somewhat disagree	18 (12.2%)	9 (10.8%)	27 (11.7%)
	Disagree	50 (33.8%)	24 (28.9%)	74 (32.0%)
	Strongly disagree	59 (39.9%)	29 (34.9%)	88 (38.1%)
How much do you agree with the following statement: 'driverless cars should be segregated and used only on dedicated roads/lanes'	Strongly agree	7 (4.7%)	5 (6.0%)	12 (5.2%)
	Agree	8 (5.4%)	9 (10.8%)	17 (7.4%)
	Somewhat agree	22 (14.9%)	17 (20.5%)	39 (16.9%)
	Neither agree nor disagree	18 (12.2%)	14 (16.9%)	32 (13.9%)
	Somewhat disagree	18 (12.2%)	13 (15.7%)	31 (13.4%)
	Disagree	41 (27.7%)	11 (13.3%)	52 (22.5%)
	Strongly disagree	34 (23.0%)	14 (16.9%)	48 (20.8%)
How much do you agree with the following statement: 'I would trust manufacturer or government assurances that driverless cars were safe'	Strongly agree	25 (16.9%)	9 (10.8%)	34 (14.7%)
	Agree	49 (33.1%)	20 (24.1%)	69 (29.9%)
	Somewhat agree	35 (23.6%)	23 (27.7%)	58 (25.1%)
	Neither agree nor disagree	13 (8.8%)	9 (10.8%)	22 (9.5%)
	Somewhat disagree	16 (10.8%)	13 (15.7%)	29 (12.6%)
	Disagree			

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
	Strongly disagree	6 (4.1%)	3 (3.6%)	9 (3.9%)
		4 (2.7%)	6 (7.2%)	10 94.3%)
How much do you agree with the following statement: 'I enjoy driving too much to ever want a driverless vehicle'	Strongly agree	0 (0.0%)	1 (1.2%)	1 (0.4%)
	Agree	5 (3.4%)	3 (3.6%)	8 (3.5%)
	Somewhat agree	10 (6.8%)	16 (19.3%)	26 (11.3%)
	Neither agree nor disagree	22 (14.9%)	16 (19.3%)	38 (16.5%)
	Somewhat disagree	14 (9.5%)	17 (20.5%)	31 (13.4%)
	Disagree	39 (26.4%)	17 (20.5%)	56 (24.2%)
	Strongly disagree	58 (39.2%)	13 (15.7%)	71 (30.7%)
How much do you agree with the following statement: 'a driverless vehicle would increase my mobility'	Strongly agree	35 (23.6%)	17 (20.7%)	52 (22.6%)
	Agree	29 (19.6%)	13 (15.9%)	42 (18.3%)
	Somewhat agree	15 (10.1%)	10 (12.2%)	25 (10.9%)
	Neither agree nor disagree	35 (23.6%)	23 (28.0%)	58 (25.2%)
	Somewhat disagree	4 (2.7%)	2 (2.4%)	6 (2.6%)
	Disagree	18 (12.2%)	14 (17.1%)	32 (13.9%)
	Strongly disagree	12 (8.1%)	3 (3.7%)	15 (6.5%)
How much do you agree with the following statement: 'a driverless vehicle would reduce the stress of driving'	Strongly agree	65 (43.9%)	21 (25.3%)	86 (37.2%
	Agree	48 (32.4%)	26 (31.3%)	74 (32.0%
	Somewhat agree	23 (15.5%)	18 (21.7%)	41 (17.7%
	Neither agree nor disagree	4 (2.7%)	9 (10.8%)	13 (5.6%
	Somewhat disagree	5 (3.4%)	7 (8.4%)	12 (5.2%
	Disagree	1 (0.7%)	1 (1.2%)	2 (0.9%
	Strongly disagree	2 (1.4%)	1 (1.2%)	3 (1.3%)

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
How much do you agree with following statement: 'My extent of understanding vehicle automation and driverless technology is a significant factor in my feelings towards driverless vehicles'	Strongly agree	43 (29.1%)	22 (26.2%)	65 (28.0%)
	Agree	51 (34.5%)	33 (39.3%)	84 (36.2%)
	Somewhat agree	31 (20.9%)	15 (17.9%)	46 (19.8%)
	Neither agree nor disagree	17 (11.5%)	9 (10.7%)	26 (11.2%)
	Somewhat disagree	2 (1.4%)	2 (2.4%)	4 (1.7%)
	Disagree	3 (2.0%)	2 (2.4%)	5 (2.2%)
	Strongly disagree	1 (0.7%)	1 (1.2%)	2 (0.9%)
How much do you agree with the following statement: 'I am suspicious of a driverless vehicles intent, action or outputs'	Strongly agree	2 (1.4%)	1 (1.2%)	3 (1.3%)
	Agree	8 (5.4%)	4 (4.8%)	12 (5.2%)
	Somewhat agree	10 (6.8%)	18 (21.7%)	28 (12.1%)
	Neither agree nor disagree	14 (9.5%)	8 (9.6%)	22 (9.5%)
	Somewhat disagree	13 (8.8%)	13 (15.7%)	26 (11.3%)
	Disagree	55 (37.2%)	27 (32.5%)	82 (35.5%)
	Strongly disagree	46 (31.1%)	12 (14.5%)	58 (25.1%)
How much do you agree with the following statement: 'I am wary of driverless vehicles'	Strongly agree	2 (1.4%)	6 (7.2%)	8 (3.5%)
	Agree	5 (3.4%)	10 (12.0%)	15 (6.5%)
	Somewhat agree	30 (20.3%)	17 (20.5%)	47 (20.3%)
	Neither agree nor disagree	12 (8.1%)	8 (9.6%)	20 (8.7%)
	Somewhat disagree	15 (10.1%)	11 (13.3%)	26 (11.3%)
	Disagree	39 (26.4%)	20 (24.1%)	59 (25.5%)
		45 (30.4%)	11 (13.3%)	56 (24.2%)

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
	Strongly disagree			
How much do you agree with the following statement: 'I am confident in a driverless vehicles performance'	Strongly agree	38 (25.7%)	10 (12.0%)	48 (20.8%)
	Agree	50 (33.8%)	22 (26.5%)	72 (31.2%)
	Somewhat agree	35 (23.6%)	24 (28.9%)	59 (25.5%)
	Neither agree nor disagree	17 (11.5%)	15 (18.1%)	32 (13.9%)
	Somewhat disagree	6 (4.1%)	6 (7.2%)	12 (5.2%)
	Disagree	1 (0.7%)	6 (7.2%)	7 (3.0%)
	Strongly disagree	1 (0.7%)	0 (0.0%)	1 (0.4%)
How much do you agree with the following statement: 'A driverless vehicle will provide safety to both the occupants of the vehicle and pedestrians'	Strongly agree	50 (33.8%)	11 (13.3%)	61 (26.4%)
	Agree	57 (38.5%)	29 (34.9%)	86 (37.2%)
	Somewhat agree	26 (17.6%)	19 (22.9%)	45 (19.5%)
	Neither agree nor disagree	10 (6.8%)	18 (21.7%)	28 (12.1%)
	Somewhat disagree	2 (1.4%)	4 (4.8%)	6 (2.6%)
	Disagree	2 (1.4%)	1 (1.2%)	3 (1.3%)
	Strongly disagree	1 (0.7%)	1 (1.2%)	2 (0.9%)
How much do you agree with the following statement: 'A driverless vehicle will be dependable in all situations'	Strongly agree	12 (8.1%)	4 (4.8%)	16 (6.9%)
	Agree	41 (27.7%)	14 (16.9%)	55 (23.8%)
	Somewhat agree	36 (24.3%)	25 (30.1%)	61 (26.4%)
	Neither agree nor disagree	22 (14.9%)	8 (9.6%)	30 (13.0%)
	Somewhat disagree	22 (14.9%)	16 (19.3%)	38 (16.5%)
	Disagree	12 (8.1%)	11 (13.3%)	23 (10.0%)
	Strongly disagree	3 (2.0%)	5 (6.0%)	8 (3.5%)
How much do you agree with	Strongly agree	33 (22.3%)	5 (6.0%)	38 (16.5%)

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
the following statement: 'A driverless vehicle will be reliable'	Agree	59 (39.9%)	30 (36.1%)	89 (38.5%)
	Somewhat agree	30 (20.3%)	16 (19.3%)	46 (19.9%)
	Neither agree nor disagree	18 (12.2%)	21 (25.3%)	39 (16.9%)
	Somewhat disagree	7 (4.7%)	7 (8.4%)	14 (6.1%)
	Disagree	0 (0.0%)	3 (3.6%)	3 (1.3%)
	Strongly disagree	1 (0.7%)	1 (1.2%)	2 (0.9%)
How much do you agree with the following statement: 'I can trust a driverless vehicle	Strongly agree	33 (22.3%)	11 (13.1%)	44 (18.9%)
	Agree	59 (39.9%)	24 (28.6%)	83 (35.8%)
	Somewhat agree	31 (20.9%)	22 (26.2%)	53 (22.8%)
	Neither agree nor disagree	18 (12.2%)	13 (15.7%)	31 (13.4%)
	Somewhat disagree	5 (3.4%)	9 (10.8%)	14 (6.0%)
	Disagree	1 (0.7%)	5 (6.0%)	6 (2.6%)
	Strongly disagree	1 (0.7%)	0 (0%)	1 (0.4%)
Which sector or area do you believe will benefit the most from driverless vehicles (please only select one)	Car industry	7 (4.8%)	8 (9.9%)	15 (6.6%)
	Businesses	19 (13.0%)	9 (11.1%)	28 (12.3%)
	Society	66 (45.2%)	23 (28.4%)	89 (39.2%)
	Environment	19 (13.0%)	12 (14.8%)	31 (13.7%)
	Individuals	24 (16.4%)	19 (23.5%)	43 (18.9%)
	Not sure	11 (7.5%)	10 (12.3%)	21 (9.3%)
Do you think driverless vehicles will bring more freedom	Definitely yes	74 (50.3%)	27 (33.3%)	101 (44.3%)
	Probably yes	45 (30.6%)	31 (38.3%)	76 (33.3%)
	Might or might not	17 (11.6%)	15 (18.5%)	32 (14.0%)
	Probably not	9 (6.1%)	8 (9.9%)	17 (7.5%)
	Definitely not	2 (1.4%)	0 (0.0%)	2 (0.9%)

Question	Response selection	GATEway respondents	Internet respondents	Total respondents
Do you think driverless vehicles will make future cities better to live and travel in?	Much better	85 (57.8%)	32 (39.5%)	117 (51.3%)
	Moderately better	28 (19.0%)	20 (24.7%)	48 (21.1%)
	Slightly better	20 (13.6%)	13 (16.0%)	33 (14.5%)
	About the same	7 (4.8%)	14 (17.3%)	21 (9.2%)
	Slightly worse	4 (2.7%)	1 (1.2%)	5 (2.2%)
	Moderately worse	2 (1.4%)	0 (0.0%)	2 (0.9%)
	Much worse	1 (0.7%)	1 (1.2%)	2 (0.9%)
Do you think driverless vehicles will 'be the norm on UK roads' within the next 5-10 years?	Definitely yes	17 (11.6%)	5 (6.2%)	22 (9.6%)
	Probably yes	44 (29.9%)	15 (18.5%)	59 (25.9%)
	Might or might not	37 (25.2%)	16 (19.8%)	53 (23.2%)
	Probably not	39 (26.5%)	38 (46.9%)	77 (33.8%)
	Definitely not	10 (6.8%)	7 (8.6%)	17 (7.5%)
Do you think driverless vehicles will 'be the norm on UK roads' within the next 15-20 years?	Definitely yes	63 (42.9%)	22 (26.8%)	85 (37.1%)
	Probably yes	59 (40.1%)	32 (39.0%)	91 (39.7%)
	Might or might not	11 (7.5%)	17 (20.7%)	28 (12.2%)
	Probably not	13 (8.8%)	9 (11.0%)	22 (9.6%)
	Definitely not	1 (0.7%)	2 (2.4%)	3 (1.3%)

Appendix C

All additional comments from respondents (verbatim).

GATEway Respondents

“Have a look at the case story's from TESLA, you will see quite a few of their "driverless cars" crashing at high speeds, and crossing red lights, breaking driving laws etc. Now ask yourself why aren't any other brands doing what TESLA are doing with their driverless cars and systems?”

“I think there's too many people who enjoy driving manually for driverless to be the norm in the short term, but the closer we get to making manually driven cars less attractive to people looking for a car, the better. Whether this is by law or by cost, will be interesting.”

None

“Success of driverless vehicles depend on the infrastructure (and security of) for success. In a closed - short distance environment it would be beneficial to all. Further testing and development is needed to change mindsets before they would be the norm on main roads.”

“Can you please push for driverless vehicles to be on the roads sooner!”

“I think the discussion on potential accidents - a car choosing to hit elderly vs. Younger people in case an accident cannot be avoided will be very interesting”

“My big concern re driverless cars is the possibility of a hack attack”

“Based on what we know about automated systems I would trust it more than a human being in a driving task. There are too many variables even in regular driving especially in unforeseen situations that a normal human can react well too. Every human has to be trained to drive to do automotive processes and machines have been proven to be able to do tasks better. Driving can be seen as a relative simple task to teach. It's the other humans on the road that will cause problems in its initiation process.”

“I think the technology will have a massive impact on society and the environment. Driverless cars can be cleaner and more fuel-efficient. They will be safer on the road because they will not take unnecessary risks, drive erratically, speed, or get tired. Travelling long distances will be less stressful and safer, especially for the old or disabled.”

“Cant wait ....”

“Interesting concept and once safety issues are firmly in public mind I can see there being a real shift towards driverless cars”

“I am wary of driverless vehicles in the short term but confident they will quickly (5-15 years) become substantially safer and more efficient than driven vehicles.”

“I am concerned about the inevitable situation where the vehicle chooses to kill its driver to save a greater number of external people in an unavoidable collision.”

“Many of these questions assume a level of knowledge and experience of driverless vehicles which the vast majority of respondents will not have - so the results of this survey can hardly be accurate or even representative. Certainly many of my answers have been pure guesswork or based on supposition. Sadly, I suspect that the aim of this survey is to tick the "public consultation" box ...”

“As a philosophy graduate with an interest in ethics, I find the moral arguments about driverless cars fascinating, but not as complex as they may first seem. I think that when viewed rationally, the main issue most people raise ('who would be responsible in a crash?') is not actually that important.”

“Morally, the primary consideration should be maximising the benefits and minimising the risks. Motor vehicles cause a gargantuan amount of death, illness and injury directly through collisions and indirectly through increasing inactivity, dominating road space and polluting.”

“As driverless cars already surpass humans in terms of pollution generated and collisions involved in, there is a very strong moral case that transition to driverless vehicles gets underway swiftly.”

“Of course, doing so in a rash and irresponsible manner would not be morally justifiable, yet delaying unnecessarily would be no more right or reasonable.”

“The 'I like to drive' argument is frankly worthy of contempt and should be immediately countered by facts about the harm human drivers cause.”

“Whilst it is assumed that driverless cars of the future will run on cleaner fuel sources, it should also be bourne in mind that much pollution comes from particulates caused by braking and general vehicle functioning. Perhaps it would be worth considering if driverless vehicles were able to minimise the volume of particulates emitted.”

“The sooner we can switch to full automation and electric vehicles, the safer and cleaner we'll be.”

“I have a motorcycle licence, a car licence, a HGV licence, a Coach Licence, a certificate to drive dangerous goods, I have been a transport supervisor, a Hospital Ambulance driver, a HGV delivery driver, and an Army driver.”

“I look forward to driverless vehicles becoming the norm on UK roads as soon as possible, because I want to cycle on London's roads safely, I want children to be able to play safely on London's roads, I'm tired of being abused by young impatient drivers when I don't race to the next set of traffic lights, and I find the time that I spend driving to be wasted time so I look forward with eager anticipation to being in the rear browsing the internet or being asleep.”

“I think the prospects of driverless cars being accepted depend on changes to traffic and road policies and strategy as much as the reliability as the cars.”

“Bring them on !”

“Driverless technology is exciting but with all new technologies, younger people will be more trusting of the technology than older people. It is this level of trust that will be the barrier to success for the adoption of driverless vehicles.”

“I would only doubt a driverless cars abilities in adverse weather conditions or unpredictable situations such as flooding, landslides etc, and would appreciate manual override to avoid fatalities.”

“The government has outlined Levels 0-5 of automated vehicle control, it might have been better if explanations outlined these to give a greater idea as to the expected levels of sophistication and asked questions relating to these levels, rather than just dividing questions between 'driver assist (driver still has responsibility)' and 'fully automated' (i.e. Level 5 automation).”

“Yes driverless vehicle may good for the mobility of people , but it's business that will be the big winner saving millions and meaning less jobs putting more people out of work”

“Driverless vehicles are only a small part of an intelligent integrated transport system. One should be very wary of seeing them as some sort of transport panacea.”

“An ideal transport future would have autonomous portions, but a like for like replacement of private human controlled cars with private autonomous cars would solve very few problems.”

“Additionally, the more significant use case is in moving goods rather than people.”

“I find it difficult to answer questions on trust and faith etc in a driverless vehicle as I (and the majority of the public in the UK) have no experience of driverless vehicles, and so not much to base an opinion on. A lot of the trust will come when manufacturers start bringing products to market. For example, consumers put their trust in various car brands based on factors such as quality, safety, reliability etc. I believe consumers will make similar opinions for driverless cars - for example Volvo

owners may be more likely to trust a Volvo-manufactured driverless car than one of a different brand."

"To the individual, there could be substantial benefits. However, at a city-wide level, if they reduced the cost and inconvenience of using cars, problems like congestion could get worse"

"I'm very keen for driverless cars to be introduced as I really dislike driving as my spatial awareness skills are not strong"

"My biggest worry is ability of criminals to hack into the systems"

"I think driverless cars are a great idea. I have concerns over security and software viruses affecting safety. And how this is implemented needs very careful thought as it could put public transport costs out of reach of people. This should not be a private only enterprise."

"I am far far more wary of the semi-autonomous vehicles than the full driverless vehicles It allows humans to take back control and have an accident. So if we had a society where all cars were automated it feels safer as they'd all be working within their programmed parameters for safety. But if it is 'some' driverless cars in amongst normal drivers speeding, drink driving, aggressive driving etc, then it feels unsafe to be in a driverless car at that point - if I do not have control of my driverless car I cannot make sure I get out of the way when the bad driver is near me and I do not trust the driverless car to be able to react correctly/quickly to all the possible bad 'human' driving possibilities that could occur. Does that make sense? Summary: All cars fully automated feels safe, half and half feels risky to me."

"No"

"I am visually impaired and no longer have a licence. I can't wait for a driverless vehicle to give me back my freedom."

"I have a good understanding of the technology as I work in the railway industry where digital signalling has meant that driverless trains are the norm for metro railways. I have seen first hand the improvements in safety the technology has delivered therefore I have no doubt it will deliver the same results for road transport. I am primarily a cyclist (and pedestrian) so I see on a daily basis the hazards in cities of human drivers - also for motorway and rural roads driver fatigue is a massive problem. For me personally the biggest benefits of driverless vehicles will be the improvement in safety. However I can see the enormous benefit this technology will have for elderly and disabled people who no longer can drive and find it difficult to use public transport/ walking or cycling as mobility options."

“I think that driverless vehicles will have a significant impact on our city's landscape. No longer do we need streets lined with parked vehicles or parking lots. Cars will move themselves out of the way to less dense areas”

“I also think that the model of ownership will change - one will rent time in a driverless vehicles rather than own one.”

“I am short sigted in my vision and i am legally not permitted to have a licence. I have answered the questions as if i was a driver. I am interested to see how the semi autonomous features of driverless cars might benefit people like me and our mobility”

Internet Respondents

“I love the idea of driverless vehicles but the concern is the early days where they will not be able to compensate for human drivers errors. Once all cars are driverless roads will be much more efficient with no more ripple effect of tiny human delays causing jams. Also there's been some concerning reports of crashes where the driverless vehicle has failed to detect a large white lorry in the lane next to it. Although I'm confident these issues will improve other time with more testing”

“My main concern is not with the driverless cars it's other road users. The period with non autonomous and autonomous cars will be strife with concern and blame. Were dedicated routes be driverless only I would be completely confident due to the removal or human error”

“I would love to take part in the trials please.”

“It is hard to form opinions on a hypothetical driverless system. Early systems will not be as robust as more mature systems in years to come. I would be cautious of early systems, but expect drivers to be fully replaced eventually. In the future you will not be able to buy a car that you can drive yourself.”

“I feel the biggest benefit will be to non drivers in rural areas who will be able to utilise these vehicles”

“Driverless vehicles will provide a great deal of independence, mobility and freedom to the disabled community if we are allowed to make use of them, especially those of us whose conditions mean we are unable to attain a driving licence for a normal car and so are completely dependant on another person being willing to drive us currently.”

“Driverless vehicles make sense on boring roads like motorways where we could get greater throughout from by using platoons but I would want to drive without assistance if I was driving on say the A830 Road to the Isles where it is all about the driving experience”

“I think it will be necessary to show that the existing infrastructure can adapt to driverless vehicles at the same time as showing the qualities of the vehicles themselves because our learned traffic responses are bound up with what we know about the existing infrastructure”

“No”

“Trust in driverless cars is dependent on performance which at this stage is unproven. The fear of 3 rd party intervention is a real risk. Potentially a great idea that will increase safety and people mobility in an ageing society.”

“I think there needs to be clear responsibility over the actions of the vehicle even if the driver is 'hands and eyes off'. I can see how a minority could benefit from it but the cost of it may be prohibitive to many especially in poorer urban areas where owning a car at all is difficult.”

“I'm for the introduction of driverless cars during rush hours as it would be feasible to me that by maximising the vehicle to vehicle interaction the traffic flow would become more efficient and reduce stress on the drivers and passengers...”

“The mix of external factors outside the cars control causes the levels of uncertainty I have. If all cars were automated then I would have stronger trust. Obviously next risk is their security and being hacked, but I think inside a city they would have a large benefit.”

“Everybody will benefit from driverless cars once they can trust them, but getting people to trust them will be the hardest part to implement them into society.”

“Can't wait to try them”

“I don't know enough about this subject to give useful answers. Good luck!”

“Sounds like the future might actually be happening. I can see huge positives with automatic cars but these will work better with more on the road and i doubt the majority will want this.”

“I hope driverless vehicles in future will all be EVs.”

“My main concerns around driverless cars is the reliability and ethics of the artificial intelligence.”

“My answers to a lot of the questions would ideally be prefaced by It depends on the safeguards in the technology, which I don't know much about.”

“A lot of the trustworthiness of the driverless vehicles would depend on independent reviews, safety ratings, and other people's experiences, as well as the trustworthiness of the manufacturer and the technology”

"I worry about reliability long term and cost to maintain. Could be great for people who would otherwise be housebound. It would take a lot for me to feel comfortable driving one and would worry that I couldn't trust it."

Appendix D

Information sheet presented to half of the participants who took part in the survey

The following information is taken from British government reports, a report by KPMG and academic research reports from UK and US universities.

Driverless Cars Currently on the Road

Google reports that collectively their driverless cars have driven over 1.5 million miles driverless and more than 500,000 miles without crashing. The first reported fatality in a semi-autonomous vehicle was recently reported by Tesla in over 130 million miles of driving. This fatality rate compares favourably to that of manual driving in the US (on average one fatality for every 94 million miles driven) and manual driving worldwide (one fatality per 60 million miles) but not to that of manual driving in the UK (one fatality per 178 million miles).

Predicted Societal Impacts

Driverless vehicles could enable more efficient use of road space, through developments such as platooning (vehicle groups travelling close together), narrowing of lanes, and reduced junction stops. This could lead to improved traffic flow, reduced congestion, improved fuel economy, reduced pollution emissions and lower costs. Changes in engine design could also reduce emissions and increase fuel efficiency. Driverless technology could enhance disabled or older people’s mobility, giving transport access to those who currently cannot drive. This could reduce both the need for motorists to chauffeur non-drivers and the use of subsidised public transport. The introduction of driverless vehicles could facilitate car sharing (vehicle rental services that substitute for personal vehicle ownership) leading to reductions in various costs associated with car ownership (e.g. insurance premiums, maintenance). Driverless technology could reduce driver stress and allow motorists to relax, socialise and work while traveling.

Predicted Economic Impacts

Driverless vehicle technologies could lead to improved productivity and increased trade for the UK, as industries capture part of a wider global market for

Intelligent Mobility estimated to be worth £900bn worldwide by 2025. This influx of new business and industry for driverless vehicles could create an additional 320,000 jobs in the UK by 2030, of which 25,000 would be in automotive manufacturing. It is estimated that driverless technology could increase GDP by 1% by 2030. Overall, the economic and social benefits of connected and driverless vehicles (i.e. fewer accidents, improved productivity and increased trade) are predicted to amount to somewhere in the region of £51 billion per year by 2030. Longer-term estimates predict that by 2040 these benefits will more than double, to £121 billion.

Predicted Safety Impacts

More than 90% of road traffic collisions are currently thought to be caused by human error. Recent collision avoidance technologies such as Electronic Stability Control and Autonomous Emergency Braking Systems have shown a more than 20% benefit in collision reduction. The adoption of semi-autonomous and driverless technology could save over 2,500 lives and prevent more than 25,000 serious accidents in the UK by 2030. This increased safety may reduce many common accident risks and therefore crash costs and reduce insurance premiums.

Predicted Risks and Costs

There may be reduced employment and business activity in sectors such as haulage and private vehicle hire, which rely heavily on people to drive vehicles. There could be additional short- to mid-term costs to consumers and governments as vehicles require additional equipment, services and maintenance, and changes are required to roadway infrastructure. There could be additional risks both anticipated (e.g. encouraging road users to take additional risks) or unforeseen (e.g. technological, economic, political) that could cause problems for us and future generations.