Article:

Merat, N orcid.org/0000-0003-4140-9948, Jamson, AH, Lai, FCH et al. (1 more author) (2012) Highly Automated Driving, Secondary Task Performance, and Driver State. Human Factors, 54 (5). pp. 762-771. ISSN 0018-7208

https://doi.org/10.1177/0018720812442087

Highly automated driving, secondary task performance and driver state

Natasha Merat

A. Hamish Jamson

Frank C.H. Lai,

Oliver M.J. Carsten

Institute for Transport Studies, University of Leeds, UK.

Précis: This study examined the effect of changes in workload on performance in highly automated and manual driving. Variations in workload were also observed using blink measures. Results showed good driver response to incidents in the highly automated condition and some predictions in workload levels by blink frequency measures.

Address correspondence to:

Natasha Merat

Institute for Transport Studies

University of Leeds, Leeds, LS2 9JT, UK

Email: n.merat@its.leeds.ac.uk.

Keywords: Blink duration, Blink frequency, Vehicle automation, and driver behavior.

Word Count: Abstract: 250, Text: 4,638, References: 28 Figures: 4. Tables: 2

Merat, N., Jamson, H., Lai, F. and Carsten, O. (2012). Highly automated driving, secondary task performance and driver state, Human Factors: The Journal of the Human Factors and Ergonomics Society October 2012 54: 762-771. doi:10.1177/0018720812442087

ABSTRACT

Objective: A driving simulator study compared the effect of changes in workload on performance in manual and highly automated driving. Changes in driver state were also observed by examining variations in blink patterns.

Background: With the addition of a greater number of Advanced Driver Assistance Systems in vehicles, the driver's role is likely to alter in the future from an operator in manual driving to a supervisor of highly automated cars. Understanding the implications of such advancements on drivers and road safety is important.

Method: Fifty participants were recruited for this study and drove the simulator in both manual and highly automated mode. As well as comparing the effect of adjustments in driving-related workload on performance, the effect of a secondary Twenty Questions Task was also investigated.

Results: In the absence of the secondary task, drivers' response to critical incidents was similar in manual and highly automated driving conditions. The worst performance was observed when drivers were required to regain control of driving in the automated mode, whilst distracted by the secondary task. Blink frequency patterns were more consistent for manual than automated driving, but were generally suppressed during conditions of high workload.

Conclusions: Highly automated driving did not have a deleterious effect on driver performance, when attention was not diverted to the distracting secondary task.

Application: As the number of systems implemented in cars increases, an understanding of the implications of such automation on drivers' situation awareness, workload and ability to remain engaged with the driving task is important.

INTRODUCTION

In a bid to increase the safety and comfort of drivers and reduce congestion, there has been a recent surge in the implementation of various Advanced Driver Assistance Systems (ADAS) by vehicle manufacturers. Examples include Adaptive Cruise Control (ACC), Intelligent Speed Adaptation/Assistance (ISA), and Lane Keeping Assistance Systems (LKAS). Recent European research projects such as CityMobil and HAVEit have studied the effect of a higher level of vehicle automation on driver behavior, where a number of such devices are combined to produce a system which can manage all vehicle maneuvers, handling longitudinal and lateral control, reducing speed to avoid collision with vehicles, changing lane and warning drivers if unable to manage emergency situations (Reiner et al., 2008, Toffetti et al., 2010). The focus of such research is somewhat different to past North American work on the Automated Highway System (AHS) which relied on infrastructural changes for its success, and was considered too expensive to attract adequate support for deployment (Ioannou, 1997). Arguably, more success has been achieved in the U.S with driverless vehicles that do not rely on dedicated infrastructure (DARPA Grand Challenge, 2010).

A distinction is made here between the control of such driverless cars relying on *full* automation and *highly automated* vehicles which still require some monitoring of the automated systems by the driver (Flemisch, et al., 2008). In essence, driving such highly automated cars alters the drivers' role from an operator to a system supervisor who is simply required to monitor the driving scene and resume control of the vehicle, for instance during emergency events, where the limits of the automated system have been surpassed.

Whilst monitoring the automation controlling a vehicle in a quiet road might not be too onerous (McKnight & McKnight, 2003), and easily achievable, even by novice drivers, a different state of affairs is envisaged when the highly automated vehicle interacts with more complex road and traffic environments, for example: to avoid collision with other non-automated vehicles, or pedestrians/cyclists performing unexpected maneuvers. Here, drivers must understand the functionalities of the automated system to ascertain if they should resume control of the driving task. Therefore, drivers need to be aware of what is taking place in the road, how surrounding traffic is behaving and whether the automated vehicle is responding effectively. Such good Situation Awareness (SA) by drivers involves the perception and comprehension of the current conditions, as well as projection of future actions (Endsley, 1995). However, previous studies, for example in aviation, suggest that automation can have a negative effect on SA, due to over-reliance on the system, reduced vigilance and lack of understanding of the system's capabilities (Endsley, 1997).

Much of our understanding of the implications of such high levels of vehicle automation on human behavior has relied upon work on pilots, nuclear plant operators and air traffic controllers (e.g. Hancock & Parasuraman, 1992), although there is now emerging work on the effect of highly automated vehicles on factors such as driver fatigue (Desmond, Hancock & Monette, 1998; Saxby, Matthews, Hitchcock, Warm, Funke & Gantzer, 2008) situation awareness (Merat & Jamson, 2009 a, b) physiological state (Rauch, Kaussner, Krüger, Boverie & Flemisch, 2009), and workload (Young & Stanton, 2002). A common theme across many of these studies is understanding drivers' interaction with ACC (Ma & Kaber, 2005; Rudin-Brown & Parker, 2004; Stanton & Young, 2005; Seppelt & Lee, 2007), although results are conflicting. For example, whilst Ma and Kaber (2005) showed reduced mental workload in the presence of ACC, and safer driving behavior with respect to speed and headway measures, Rudin-Brown & Parker (2004) suggest that such reduced workload was likely to lead to dangerous underload, creating unsafe driving behavior when participants were required to respond to a hazard.

One area which merits better understanding in this domain is how highly automated vehicles affect driver workload, and how performance changes when periods of underload are followed by sudden and high overload. Although the effect of changes in workload on *manual* driving has been extensively studied, less is known about the effect of such transitions in workload during highly automated driving. We know from studies on automation in other domains that it is inaccurate to assume a direct reduction in workload by automation, and that automation is in fact likely to lead to a "redistribution of workload" during different stages of a task (Sarter, Woods, and Billings, 1994). For example, studies from aviation suggest that whilst automation can support pilots during situations of low workload, it may actually be problematic by failing to provide the right support during conditions of higher workload, at "time-critical, highly dynamic" phases of a flight (Sarter et al., 1994). Using the "Malleable Attentional Resources Theory" (MART), Young and Stanton (2002) propose that performance degrades in such situations because automating a task causes a temporary reduction in attentional resources, which are not as necessary when automation is in charge. Such reduction in resources is then detrimental to performance if workload suddenly intensifies, for example due to unexpected changes in the driving environment.

A variety of objective and subjective tools have been used to observe the effect of changes in workload on operator state and performance. One attractive methodology, which provides a timely and objective understanding, is measures of drivers' psychophysiological state. Examples of such measures include: (i) Evoke Related Potentials (ii) eye-tracking measures: frequency, interval and duration of blinks and (iii) heart-related measures: heart rate and heart rate variability (Wilson & Russell, 2003). Due to their ease of use and less intrusive nature, heart- and eye-related measures are used extensively for studying physiological response to operator workload, fatigue, and stress (e.g. Byrne & Parasuraman, 1996; de Waard, 1996; Ryu & Myung, 2005; Recarte, Pérez, Conchillo & Nunes, 2008; see also Neumann & Lipp, 2002, for a review). Unfortunately, findings from such psychophysiological measures are somewhat conflicting. For example, workload has been reported to increase (de Waard, 1996) and decrease (Brooking, Wilson & Swain, 1996; Van Orden, Limbert, Makeig & Jung, 2001) eye blink frequency. Although there is some argument that such eye-tracking measures should only be used to estimate the effect of *visual* and not auditory workload on drivers (de Waard, 1996), the continued use of various sound-based infotainment systems in the car increases the need to understand the effect of *non-visual* secondary tasks on driver performance and workload. One recent account for the mixed results on the relationship between workload and blinks is provided by Recarte et al. (2008) who distinguish between the effect of the visual aspects of driving on eye-blinks, and that of cognitive (non-visual) aspects of the driving task. These authors speculate that blink rates can distinguish between visual and mental workload, and that an increase in visual workload (e.g. by driving or a visual search task) leads to blink inhibition. Participants are thought to suppress blinks to help decision making and response to a task. However, because such blink inhibitory mechanisms rely on attentional resources, blink rate increases when attention during a visual task is taken away by a demanding cognitive task.

Objectives

The aim of this study was to examine the effect of driving a highly automated vehicle on driver behavior, and observe how transitions from quite low to unexpectedly high levels of workload influence driving performance and physiological state. Changes in workload were manipulated by: (i) using a non-driving related in-vehicle secondary task (Twenty Questions Task - TQT) and (ii) choreographing an incident in the lane occupied by the driver, encouraging participants to change lane at designated points. The effect of vehicle automation and workload on drivers' awareness of the driving environment was also assessed by comparing their response to the critical incident in manual and automated driving.

METHOD

Participants

Fifty participants were recruited for this study. They ranged in age from 28 to 68 years (Mean = 47.38, SD = 10.37). Participants had more than 10 years' driving experience, and drove 24,370 miles a year on average (SD =17,732 miles). Following ethical approval from the University of Leeds, participants were recruited using a newspaper advertisement, and were paid for taking part in the study.

Equipment and the automated controllers

The University of Leeds Driving Simulator (UoLDS), incorporating an eight-degree-of-freedom motion system, was used for this study. The vehicle cab, based around a Jaguar S-type vehicle has all driver controls fully operational and is housed within a 4-m diameter spherical projection dome. The front road scene encompasses a horizontal field of view of 250^{o}, and three rear projectors display the scenes in the rear view and side mirrors. The simulator is also equipped with v4.5 of the Seeing Machines faceLAB eye-tracker, with its cameras mounted on the vehicle dashboard (Figure 1).

Figure 1 about here

To transfer from manual to highly automated driving, participants used a button on the steering wheel to engage the automation. Disengagement of automation was possible by pressing the same button, or by turning the steering wheel more than 3^o, or by pressing the brake pedal. Pressing the button engaged both the longitudinal and lateral controller of the vehicle at the same time. Design of the longitudinal controller was based around an ACC with a fixed desired target speed of 70 mph and a set headway of 1.5 seconds, although drivers were able to adjust this speed once at the start of the experiment. The lateral controller design was similar to a Lane Keeping Assistant System (LKAS) and kept the vehicle in the centre of whichever lane they were driving in at the time of its engagement.

Experimental Design

A fully within subjects, 3-factor repeated measures design was used to study the effect of vehicle automation and secondary task performance on driving performance and physiological state. The three independent variables used in this study each had two levels as follows: Drive (manual, automated), Secondary task performance (TQT, No TQT) and Scenario (Critical Incident, No Critical Incident). Therefore, for all participants the manual and automated drives consisted of (i) sections of road with ‘free’ driving: no TQT and no critical incident, (ii) driving-related workload imposed by the critical incident, (iii) non-driving related workload imposed by the TQT and (iv) high workload where they were required to negotiate the incident whilst also performing the TQT. Therefore, at least for the manual driving condition, we can suggest a hypothetical variation in workload, as shown in (Table 1).

Table 1 about here

Procedure

All drivers were provided with a brief description of the experiment's rationale and following consent to take part in the study, they were briefed on the safety and evacuation procedures of the simulator. A familiarization stage followed, whereby participants had the opportunity to test the functionality of the controllers. Participants were also briefed on the TQT and practiced this task whilst seated in the stationary simulator. They then drove the practice session, driving in both manual and automated mode, learning how to switch between the two modes and practicing the driving and TQT together. The researcher was present throughout this phase, to answer any questions. This practice session lasted around 45 minutes or until drivers were satisfied with the workings of the simulator and its automated system. Drivers then drove the simulator unaccompanied for the experiment session, with observation by the researcher from the control room. Both practice and experiment roads consisted of a simulated 3-lane U.K. motorway scene, with free flowing traffic conditions.

The experiment itself involved two drives (manual and automated) lasting around 45 minutes each, and to alleviate symptoms of fatigue, participants were granted with a short break after each drive. The order of drives was counterbalanced across participants, with half of the participants having manual control of the car before driving the automated vehicle, whilst the order of these drives was switched for the other half of participants. For the manual drive, participants were instructed to observe the posted speed limit of the road.

During the automated drive, participants were encouraged to hand control over to the vehicle as soon as the drive began, although they were reminded by the experimenter if this did not happen after around 1 minute. Upon pressing the button on the steering wheel, a small LCD panel, integrated in the speedometer, lit up to indicate activation of the system. Participants were also reminded that the controllers were comfort devices that could only manage gentle maneuvers, and were not designed to respond to critical and unexpected incidents.

The TQT was a guessing task which lasted 3 minutes and was performed twice in each drive: once in combination with the critical incident and once in a similar section of road without an incident. Based on a children's parlor game, participants are required to guess an item from within an over-riding category by asking a maximum of twenty questions. The response to each question is always only 'yes' or 'no'. This task is thought to utilize cognitive resources such as problem solving, planning and working memory, whilst its 'question and answer' format is similar to a telephone conversation, where performance is quantifiable (Horrey, Lesch & Garabet, 2009). The TQT started automatically at a designated point within the drive. A pre- recorded message was played via the vehicle's speakers, announcing that the item to be guessed belonged to one of two categories ('fruit and vegetable' or 'animal'). The driver then attempted to guess the identity of the item with a series of questions. The experimenter used a remote keyboard to provide the 'yes' or 'no' answers to drivers. Drivers were allowed to ask up to twenty guesses per item and were permitted to give up and move to the next question, if they so wished. The number of guesses and correct answers were recorded for each subject. The items chosen for this task were selected from a pre-piloted list provided by Horrey (personal communication, 2009, see Horrey et al., 2009), although some of the terms were Anglicized to ensure comprehension by our British participants.

The critical incidents involved an obstruction in the driver's lane, and drivers were pre-warned about the approaching incident via a Variable Message Sign (VMS) which was placed approximately 1500 meters before the Incident. Drivers were required to react to the incident by changing lane, or risk being stranded in their lane due to maneuvers of the surrounding traffic. This meant that in the automated drive, drivers were required to take control back from the automated system, using any one of the three methods described above.

RESULTS

Three sets of analyses are reported here. First, the effect of vehicle automation, secondary task performance and critical scenarios on vehicle-related metrics and driver behavior was examined. Then, the effect of the above variables on drivers' blink rate and duration was observed. Finally, drivers' performance in the TQT was analyzed. In each case, unless otherwise stated, results were analyzed using a repeated measures analysis of variance (ANOVA), where violations of sphericity were detected using Mauchly's Test. Post hoc Bonferroni tests were used for paired comparisons. A summary of the main results is provided in Table 2.

Table 2 about here

Driving performance measures

Average speed was lower (67.62 mph) in the highly automated drive, compared to when participants drove manually (70.46 mph) (F (1, 49) = 24.16, $p < .0001$, $\eta^2 = 0.33$). There was also a significant effect of the TQT on mean speed (F (1, 49) = 21.23, $p < .0001$, $\eta^2 = 0.30$), with participants driving slower in the presence of the TQT (68.53 mph vs. 69.56 mph). Naturally, average speed was lower during the critical incident scenarios (F (1, 49) = 688.33, $p < .0001$, $\eta^2 = 0.93$). The interactions between TQT and Scenario, and TQT, Drive and Scenario were both found to be significant (F (1, 49) = 8.35, $p < .01$, $\eta^2 = 0.15$ and F (1, 49) = 7.50, $p < .01$, $\eta^2 = 0.13$ respectively). As shown in Figure 2, during automated driving, participants reduced their speed by around the same level in response to the CI scenarios, regardless of whether or not they were performing the TQT. However, they reacted more effectively during manual driving, reducing their speed to a greater extent during the TQT, compared to when TQT was absent.

Figure 2 about here

To compare participants' response time to the critical incident scenarios in the manual and automated driving conditions, their lane change maneuvers were assessed, upon approach to the VMS. Here, the time at which drivers performed their final lane change, as advised by the VMS, was calculated. The position of the VMS was used as the starting point for this calculation, with any negative values denoting lane changes before the VMS and any positive values suggesting response to the VMS. It was also assumed that only lane changes completed between encountering the VMS sign and arriving at the first traffic cone (denoting the start of the critical incident) were those strategically initiated by the driver, as the traffic cones forced drivers to change lane.

Results showed that in the absence of the TQT, the proportion of drivers changing lane between the VMS and the first traffic cone in the manual and automated driving conditions was similar (83 % vs. 81%), whilst when performing the TQT, more drivers changed lane in the automated drive [71% vs. 64%. (X^2 (1, N=50) = 21.37, $p < .0001$)]. The time taken to change lane averaged around 30 seconds, regardless of driving condition or TQT presence.

Performance on the Twenty Questions Task

The average number of questions and correct answers to the TQT were calculated for each drive, and response during the critical incident scenarios was compared by conducting a 2 Drive (manual, automated) x 2 Scenario (No CI, CI) ANOVA. Results showed a significant effect of Scenario on the number of questions (F (1, 49) = 6.59, $p < .05$, $\eta^2 = .12$) with more questions asked in the absence of the critical incident (14.76 vs. 13.82). Results also showed an interaction between Drive and Scenario, with more questions in the automated drive than the manual drive when there was no incident, whilst the reverse pattern was observed during the critical incident, with marginally more questions asked during the manual drive. There was no difference in the number of correct responses, with drivers averaging around .97 correct answers (ranging between 1 and 4), regardless of drive type and scenario.

Driver state measures

FaceLAB v4.5 was used for eye-tracking. The mean number of blinks per second and the mean length of time the eyes were closed during each blink over a 60 second window were measured throughout the drive. As the quality of eye-tracking data was poor for 6 of the participants, with many missing values, results are reported for the remaining 44. Analyses showed a main effect of TQT and Scenario on the average frequency of blinks. Blink frequency was found to be higher in the presence of the TQT (0.49 Hz with TQT vs. 0.46 Hz with No TQT), ($F (1, 43) = 9.17$, $p < .01$, $\eta^2 = .18$), but a lower blink frequency was observed overall when drivers were required to negotiate the critical incident (0.47 Hz during the CI vs. 0.48 Hz. in the absence of the CI) ($F (1, 43) = 9.04$, $p < .01$, $\eta^2 = .17$). Mean blink frequency was not found to be different in the manual and automated driving conditions. However, there were highly significant interactions between the TQT and Drive ($F (1, 43) = 42.25$, $p < .0001$, $\eta^2 = .49$), and the TQT and Scenario ($F (1, 43) = 15.67$, $p < .0001$, $\eta^2 = .27$), and a three way interaction between TQT, Drive and Scenario ($F (1, 43) = 17.40$, $p < .0001$, $\eta^2 = .29$). As shown in Figure 3, in the absence of the TQT, blink frequency was higher during the automated drive, and a further increase was seen when participants were faced with the critical incident. However, in the presence of the TQT, whilst blink frequency remained high during the automated drive in the 'free' driving condition (No CI) the pattern reversed during the CI, where blink frequency was lower than that for the manual drive.

Figure 3 about here

For blink duration, only a main effect of Scenario was observed $F (1, 43) = 4.42$, $p < .05$, $\eta^2 = .09$), with marginally longer blink durations in the absence of the critical incident (0.179 seconds versus .176 seconds). There was a significant interaction between TQT and Scenario ($F (1, 43) = 13.80$, $p < .001$, $\eta^2 = .24$) and a three way interaction between TQT, Scenario and Drive ($F (1, 43) = 26.10$, $p < .0001$, $\eta^2 = .38$). In the absence of the TQT, blink duration was longer for the automated drive, only when participants were required to negotiate the critical incident. When drivers were required to perform the TQT, the same pattern was seen for blink duration in the absence of the CI. However, if they were required to negotiate the critical incident at the same time as performing the TQT, the reverse pattern was observed for blink duration, where longer durations of blink were seen for the manual drive (0.178 seconds), compared to the automated drive (0.173 seconds, Figure 4).

Figure 4 about here

DISCUSSION

The main objective of this study was to examine the effect of driving a highly automated vehicle on driver behavior, investigating how changes in workload affected driver performance. In addition to increasing driving-related workload by adding a critical incident in the road, supplementary workload was imposed on drivers by means of a distracting secondary task. Young and Stanton's Malleable Attentional Resources Theory (2002) was used to examine how changes from relatively low to unexpectedly high workload during automated driving affected drivers' control of the vehicle in response to a critical incident. As well as observing driving-related measures, we examined physiological changes, by measuring drivers' blink behavior. Here, we considered the theory put forward by Recarte et al (2008), which suggests that whilst a demanding visual scene, such as the CI used in our study, can lead to blink inhibitions (to enhance drivers' perception of the scene) the addition of a demanding non-visual (cognitive) task, removes such inhibitory mechanisms and leads to an increase in blink frequency.

Our study found that, in the absence of the TQT, drivers reduced their speed in response to the critical incident during both the manual and highly automated drives.

Participants were therefore able to comprehend forthcoming changes in the driving environment and respond effectively, showing good awareness of their surrounding environment during the highly automated drive. However, when required to perform the TQT, drivers slowed down more in the manual than the highly automated drive, when faced with the CI. Therefore, they might have been aware of their resource limitations during the manual drive and reduced their driving speed to a level suitable for managing the incident and performing the TQT. However, they were not able to reduce speed as adequately during the automated drive, perhaps because their attentional resources were directed away from the automated driving task and mainly engaged in the TQT. Under these conditions, performance was found to deteriorate, when the relatively manageable workload required for automated driving and the TQT was suddenly and unexpectedly elevated during negotiation of the CI, prohibiting drivers from responding appropriately and in time, when required to regain control of the driving task (Young & Stanton, 2002).

When performance of the secondary TQT was not required, the proportion of drivers changing lane in response to the VMS messages was found to be similar for the manual and highly automated drives and the time taken to change lane was also comparable. Therefore, driving in the automated mode without the TQT distraction did not affect comprehension of the VMS advice or their understanding of the automated system’s capabilities. However, when drivers were required to engage in the TQT, fewer lane changes were made, overall, and especially during the manual driving condition. When compared with the results on average speed above, these results suggest that whilst participants were able to promptly change lane in the automated drive, they were unable to follow this maneuver with a timely reduction in speed.

A tradeoff between primary and secondary task performance was observed during performance of the TQT, where drivers asked more questions during the ‘free’ driving (no CI) section of the automated drive, presumably because the control of driving by the system reduced their overall workload. However, as circumstances changed during the CI scenario and drivers’ full attention was needed to regain control of the driving task, to change lane/slow down, fewer resources were then dedicated to the TQT, and a fall in the number of questions asked was observed. Interestingly, such variations in workload did not have an effect on the number of correct answers.

The changes in workload imposed by the CI and TQT task showed some interesting interactions for blink frequency and duration. Overall, our results concurred with the arguments proposed by Recarte et al. (2008). For the manual drive, when TQT was absent, increased complexity of the visual scene (from no CI to CI) lead to a small fall in blink rate and duration, as participants attempted to obtain as much visual information as possible during the CI. However, adding the TQT increased blink rate during the manual drive for both the No CI and CI conditions, concurring with proposals that blink inhibition is suppressed when drivers’ attention is directed towards an attention demanding non-visual task (Recarte et al., 2008). A more complex relationship is seen for blink frequency during the automated drive, and changes in this measure provide some insight into how sudden and unexpected alterations in workload might affect this physiological state, and also where drivers’ attention is directed during such transitions from low to high workload between driving and the TQT. In the absence of the TQT and CI, the rate and duration of blinks for automated driving was very similar to that of manual driving. However, when faced with the critical incident in the automated drive, rate of blinks was substantially higher than manual driving. This may well be a feature of blink patterns in automated driving, where participants do not attend as much to the driving environment, as this is dealt with by the automated system. On the other hand, when engagement in the TQT was required, whilst the pattern of blinks was again similar for automated and manual driving in the no CI condition, an interaction was observed during the CI, where by far the lowest frequency of blinks was seen. Here, blink suppression was at its highest, since a sudden change in demand from the driving task and an unexpected need to regain control provoked drivers to obtain as much information as possible about the visual scene, where moments before their attention may have been directed mostly towards the TQT.

To conclude, this study showed that when drivers' attention was not diverted towards a demanding secondary task, performance in highly automated driving was similar to that of manual driving. Overall, the worst performance was observed when drivers in the automated mode were required to regain control of driving, whilst distracted by the secondary task. Ensuring the driver is kept in the loop is therefore vital, as one attraction of highly automated driving is the freedom to engage in other non-driving-based tasks. In terms of using driver state measures, some promising results were observed for blink frequency, but further studies using a combination of subjective and objective measures are required to understand the changes in workload experienced by participants in such highly automated driving conditions. Whilst the caveat for using such driver state measures in complex workload conditions is acknowledged (Brookings et al., 1996), there is an advantage to using such non-intrusive tools for real-time observation of driver workload, for example in order to warn drivers of dangerous underload/overload situations.

REFERENCES

Baddeley, A.D. (2000). The episodic buffer: a new component of working memory? Trends in Cognitive Sciences, Volume 4, Issue 11, 1 November 2000, Pages 417-423

Bainbridge, L. (1983). The ironies of automation. *Automatica,* 19., 775-779.

Brand, J.L. (1998). Driver Out of the Loop? *Ergonomics in Design: The Quarterly of Human Factors Applications*, 6, 26-31.

Brookings, J.B., Wilson, G.F. & Swain, C.R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology,* 42(3), 361-377.

Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology*, 42, 249-268.

Cowan, N. (1995). *Attention and memory: An integrated frame-work*. New York: Oxford University Press.

DARPA (2010). DARPA Grand Challenge Winner: Stanley the Robot. *Popular Mechanics.* http://www.popularmechanics.com/technology/engineering/robots/2169012. Retrieved 05 January 2011.

Desmond, P., Hancock, P. & Monette, J. (1998). Fatigue and Automation-Induced Impairments in Simulated Driving Perform ance, *Journal Transportation Research Record: Journal of the Transportation Research Board*, 1628, 8-14.

de Waard, D. (1996). *The measurement of drivers' mental workload*. PhD t

hesis. University of Groningen, Traffic Research Centre.

deWaard D. & Brookhuis, K. (1997). On the measurement of driver mental workload. In Rothengatter, J.A, & Carbonell, V. E. (Eds). *Traffic and Transport Psychology*, Pergamon, Amsterdam, 161-171.

Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32-64.

Endsley, M. (1997). Situation awareness, automation and free flight. Presentation at the FAA/Eurocontrol Air Traffic Management R&D Semina, Saclay, France.

Endsley, M. & Kaber, D. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics,* 42(3), 462-492.

Flemisch, F., Kelsch, J., Löper, C., Schieben, A., & Schindler, J. (2008). Automation spectrum, inner/outer compatibility and other potentially useful human factors concepts for assistance and automation. In D. de Waard, G.R.J. Hockey, P.Nickel, & K.A. Brookhuis (Eds). *Human Factors Issues in Complex System Performance* (pp. 257-272). Maastricht, The Netherlands: Shaker Publishing.

Hancock, P.A. & Parasuraman, R. (1992). Human Factors and safety in the design of Intelligent Vehicle-Highway Systems (IVHS). *Journal of Safety Research*, 23, 181-198.

Horrey, W.J. Lesch, M.F. & Garabet, A. (2009). Dissociation between driving performance and drivers‟ subjective estimates of performance and workload in dual-task conditions. *Journal of Safety Research*, 40, 7-12.

Ioannou, P. A. (1997) (Ed.), *Automated HighwaySystems*. Norwell, MA: Kluwer.

Kaber, D. B. & Endsley, M. R. (1997). Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety, *Process Safety Progress*, 16, 126-131.

Ma, R. & Kaber, D.B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. International Journal of Industrial *Ergonomics*, 35, 939–953.

McKnight, A.J. & McKnight, A.S. (2003). Motor vehicle technology: automation of driving tasks. Retrieved from http://www.isd.mel.nist.gov/research_areas/research_engineering/Performance_Metrics/PerMIS_2003/Proceedings/McKnight.pdf

Merat, N & Jamson, H. (2009a). How do drivers behave in a highly automated car? *Proceedings of the Fifth International Driving Symposium on Human Factor in Driver Assessment, Training and Vehicle Design* (pp. 514-521), Montana, USA.

Merat, N. & Jamson, A.H. (2009b). Is Drivers' Situation Awareness Influenced by a Fully Automated Driving Scenario? In D. de Waard, J. Godthelp, F.L. Kooi, & K.A. Brookhuis (Eds.) *Human Factors, Security and Safety* (pp. 1 - 11). Maastricht, the Netherlands: Shaker Publishing.

Meyer, D.E. & Keiras, D.E. (1997). A computational theory of executive cognitive processes and multiple-task performance. *Psychological Reviews,* 104 (1), 3-65.

Neumann, D.L. & Lipp, O.V. (2002). Spontaneous and reflexive eye activity measures of mental workload, *Australian Journal of Psychology*, 54:174- 179.

Rauch, N., Kaussner, A., Krüger, H.P., Boverie, S & Flemisch, F. (2009). The importance of driver state assessment within highly automated vehicles. *Conference Proceedings of the 16th World Congress on ITS*, Stockholm, Sweden.

Reiner, H., Amditis, A., Kunert, M., Hoess, A., Flemish, F., Krueger, H.P. Bartels, A., Beutner, A. & Pagle, K. (2008). Highly automated vehicles for intelligent transport: HAVEit approach. *Conference Proceedings of the 15th World Congress on ITS*, New York City, U.S.A.

Rudin-Brown, C.M. & Parker, H. (2004) Behavioural adaptation to adaptive cruise control (ACC): implications for preventive strategies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 7, 59-76.

Sarter, N.B. & Woods, D.D. (1995). “Strong, Silent, and ‘Out-Of-The-Loop’”. *CSEL Report* 95-TR-01.

Seppelt, B.D. & Lee, J.D. (2007). Making adaptive cruise control (ACC) limits visible. *International Journal of Human Computer Studies*, 65, 192-205.

Ryu, K. & Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics,* 35, 991-1009.

Recarte, M.A., Pérez, E., Conchillo, A.& Nunes, L.M. (2008). Mental workload and visual impairment: Differences between pupil, blink and subjective rating. *The Spanish Journal of Psychology,* 11, 374-385.

Sarter, N.B., Woods, D.D. & Billings, C.E. (1994). Automation Surprises. In Salvendy, G. (Ed) *Handbook of Human Factors and Ergonomics,* second edition, Wiley, UK.

Saxby, D.J., Matthews, G., Hitchcock, E.M., Warm, J.S., Funke, G.J. & Gantzer, T. (2008). Effects of active and passive fatigue on performance using a driving simulator. *Proceedings of the Human Factors and Ergonomics Society,* 52, 1252-1256.

Stanton, N. & Young, M.S. (2005). Driver behavior with adaptive cruise control *Ergonomics,* 48, 1294-1313.

Toffetti, A., Wilschut, E., Martens, M.H., Schieben, A., Rambaldini, A., Merat, N. & Flemisch, F. (2009). CITYMOBIL: Human Factors Issues Regarding Highly-Automated Vehicles on an eLane. Transportation Research Record: Journal of the Transportation Research Board, No. 2110, pp. 1–8. http://trb.metapress.com/content/k326q22n7635/

Wilson, G.F. & Russell, C.A. (2003). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors,* 45, 635-643.

Van Orden, K.F., Limbert, W., Makeig, S. & Jung, T.P. (2001). Eye Activity Correlates of Workload during a Visuospatial Memory Task. *Human Factors,* 43(1), 111-121.

Young, M.S. & Stanton, N.A. (2002). Malleable attentional resources theory: A new explanation for the effect of mental underload on performance. *Human Factors*, 44(3), 365-375.

AUTHOR BIOGRAPHIES

Natasha Merat is a Senior Research Fellow at the Institute for Transport Studies, University of Leeds, U.K. She was awarded her PhD in Psychology from the University of Leeds in 1999.

A. Hamish Jamson is a Principal Research Fellow at the Institute for Transport Studies, University of Leeds, U.K. and manager of the University of Leeds Driving Simulator. He very recently obtained his PhD in Motion Cueing in Driving Simulators from the University of Leeds.

Frank C.H. Lai is a Senior Research Fellow at the Institute for Transport Studies, University of Leeds, UK. He was awarded his PhD in Transport Safety from the University of Leeds in 2005.

Oliver M. J. Carsten is Professor of Transport Safety at the Institute for Transport Studies, University of Leeds, U.K. He was awarded his PhD in History from the University of Michigan in 1981.

KEY POINTS

- Studies argue that changing the driver's role from controller to supervisor may have an effect on performance and situation awareness.
- The interaction between changes in workload and highly automated driving is currently not well understood.
- This study showed good performance by drivers in the highly automated condition, in the absence of the secondary task.

- Performance in the driving and secondary task was found to be most impaired when the two were required together, and especially when drivers had to resume control, after a period of underload imposed by vehicle automation.

ACKNOWLEDGMENTS

The authors would like to thank the Engineering and Physical Sciences Research Council (EPSRC) for funding this research. They are also grateful for the continued support of Anthony Horrobin and Michael Daly for the driving simulator studies. The comments and feedback of Professor John Lee and three anonymous reviewers is also much appreciated.

Figure captions

Figure 5. Exterior view of the University of Leeds Driving Simulator and the vehicle cab, showing the faceLAB cameras and the controller button.

Figure 6. Changes in mean speed in the two drives as a result of the scenarios and the TQT (CI = Critical Incident) Error bars are standard errors.

Figure 7. Changes in blink frequency in the two drives as a result of the Critical Incident Scenario and TQT. Error bars are standard errors.

Figure 8. Changes in blink duration in the two Drives as a result of the Critical Incident Scenario and TQT. Error bars are standard errors.

Table Captions

Table 1 – Hypothetical changes in workload by task, for the manual driving condition.

Table 2 – Summary of all main effects (standard errors in brackets)

Table 1

Workload	**Task**	**Likely demand on attentional resources**
Low	'Free driving'	Little demand on resources, no interaction with other traffic.
Medium	Critical incident	Higher demand on resources due to a need to attend to the VMS message, followed by deciding and acting upon a lane change.
	'Free driving' + TQT	Less demand from the driving task, but more resources likely to be required for to the TQT.
High	Critical incident plus TQT	Highest demand as attentional resources are divided between negotiating the critical incident as well as responding to the TQT.

Table 2

	Automation		**TQT**		**Critical incident**	
	Present (ACC)	***Absent (Manual)***	***Present***	***Absent***	***Present***	***Absent***
Speed (mph)	67.64 (0.73)	70.46 (0.83)	69.56 (0.74)	68.53 (0.72)	65.15 (0.70)	72.94 (0.77)
TQT performance	-	-	-	-	13.82 (0.48)	14.760 (0.52)
Mean Blink Frequency	-	-	0.49 (0.03)	0.46 (0.02)	0.47 (0.03)	0.48 (0.02)
Mean Blink Duration	-	-	-	-	0.18 (0.0)	0.18 (0.0)

FIGURE 1

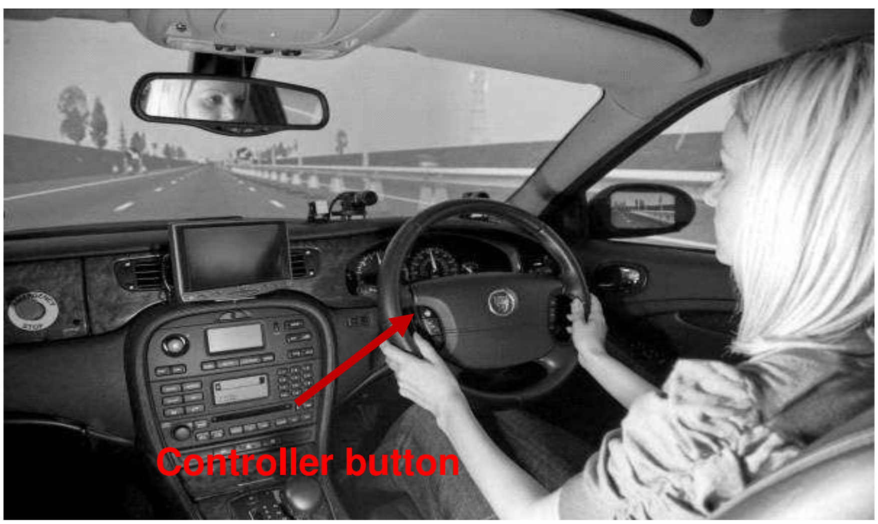


FIGURE 2

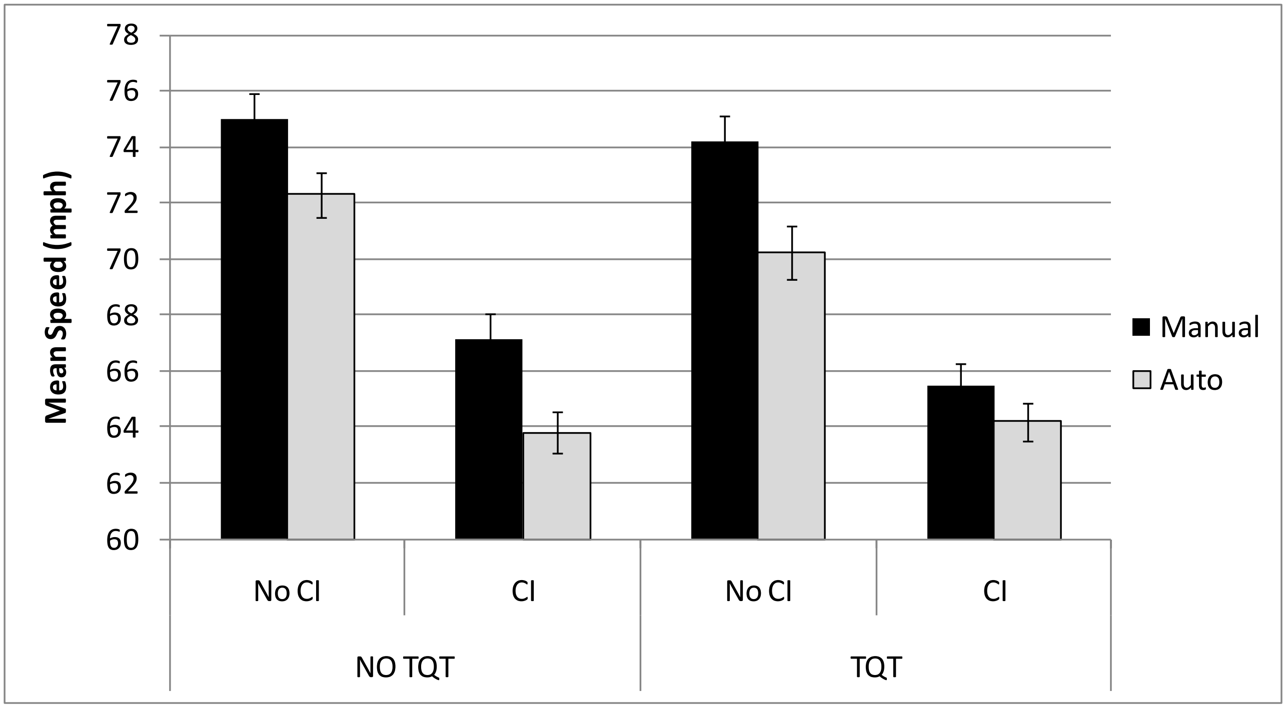

78
76
74
72
70
68
66
64
62
60
Mean Speed (mph)
Manual
Auto
No CI
CI
No CI
CI
NO TQT
TQT

FIGURE 3

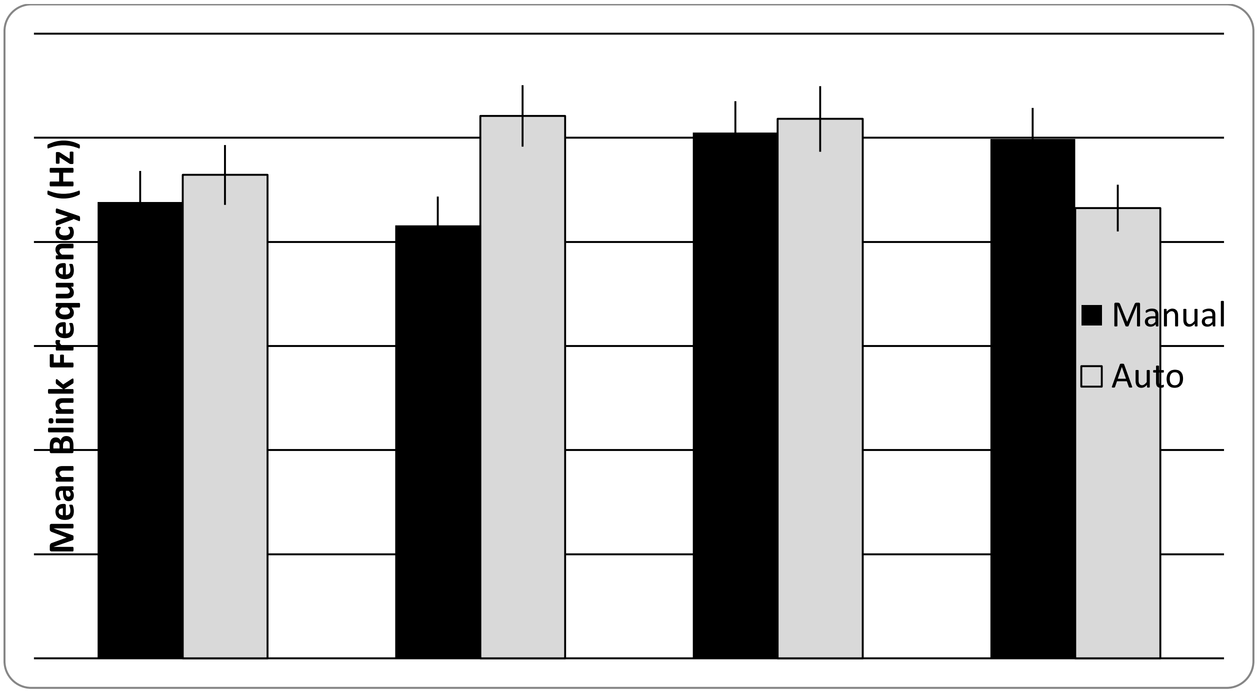
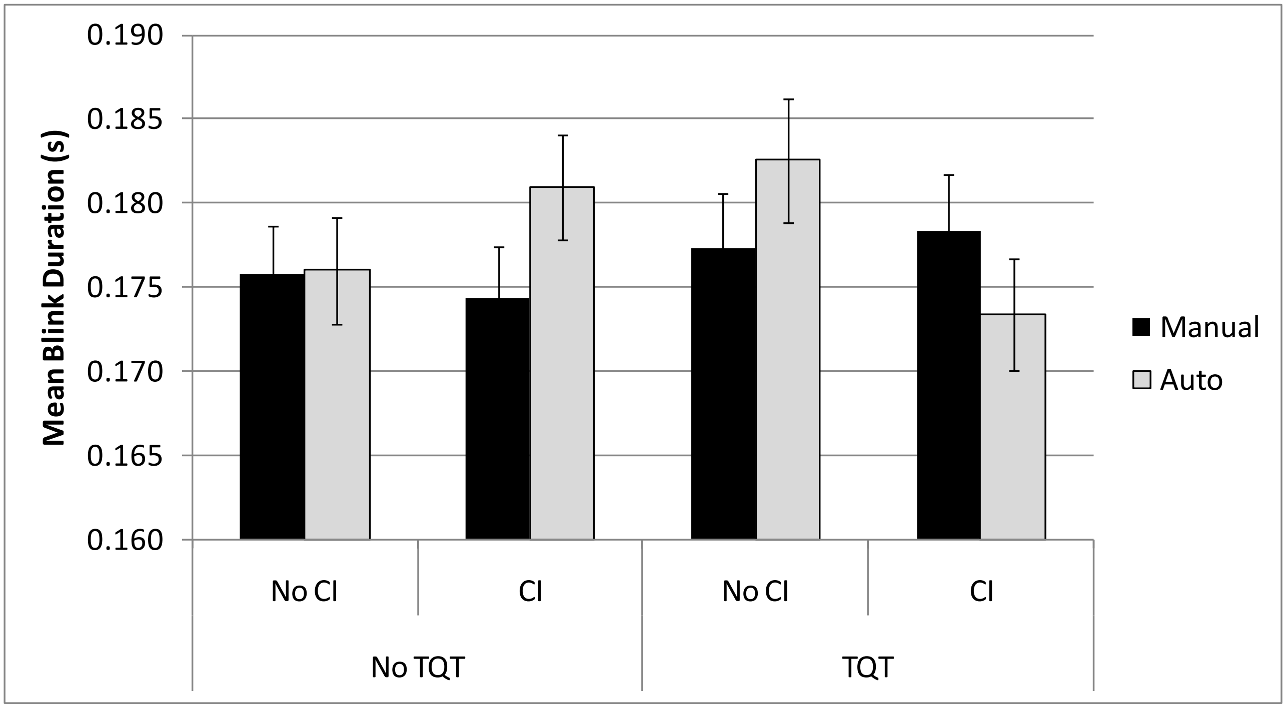


FIGURE 4


0.190
0.185
0.180
0.175
0.170
0.165
0.160
Mean Blink Duration (s)
Manual
Auto
No CI
CI
No CI
CI
No TQT
TQT