Transition to Manual: comparing simulator with on-road control transitions

Eriksson A[1], Banks V A[2], Stanton N A[1]

[1]Transportation Research Group, Faculty of Engineering and the Environment, University of Southampton, Boldrewood campus, SO16 7QF, UK

[2]Human Factors Research Group, University of Nottingham, UK

Corresponding author: Alexander Eriksson Transportation Research Group, Faculty of Engineering and the Environment, University of Southampton, Boldrewood campus, SO16 7QF, UK. Email: Alexander.eriksson@soton.ac.uk

Keywords: Automated Driving; Simulator Validity; Transfer of control; Vehicle Automation; Driver Behaviour

Cite as: Eriksson, A, Banks, V. A. & Stanton, N. A. (2017). "Transition to Manual: comparing simulator with on-road control transitions." Accident Analysis & Prevention **102C**: 227-234.

Abstract

Background: Whilst previous research has explored how driver behaviour in simulators may transfer to the open road, there has been relatively little research showing the same transfer within the field of driving automation. As a consequence, most research into human-automation interaction has primarily been carried out in a research laboratory or on closed-circuit test tracks.

Objective: The aim of this study was to assess whether research into non-critical control transactions in highly automated vehicles performed in driving simulators correlate with road driving conditions.

Method: Twenty six drivers drove a highway scenario using an automated driving mode in the simulator and twelve drivers drove on a public motorway in a Tesla Model S with the Autopilot activated. Drivers were asked to relinquish, or resume control from the automation when prompted by the vehicle interface in both the simulator and on road condition.

Results: Drivers were generally faster to resume control in the on-road driving condition. However, strong positive correlations were found between the simulator and on road driving conditions for drivers transferring control to and from automation. No significant differences were found with regard to workload, perceived usefulness and satisfaction between the simulator and on-road drives.

Conclusion: The results indicate high levels of relative validity of driving simulators as a research tool for automated driving research.

Highlights

- Control transitions in automated on road and simulated driving were assessed
- Correlation analysis was carried out for control transitions on road, and in simulators.
- No differences in self-reported workload, or technology acceptance were found between conditions.
- Results indicate high levels of relative validity for the use of simulators in control transition research.

1. Introduction

Self-driving vehicles have gone from a futuristic dream to an engineering reality (Stanton, 2015), fuelled by Moore's law (Moore, 1965). Continued development of ADAS systems such as Anti-lock Braking, Automatic Emergency Brake (Banks & Stanton, 2017), Adaptive Cruise Control (Larsson et al., 2014; Seppelt & Lee, 2007; Stanton & Young, 2005; Young & Stanton, 2007), and Lane Keeping Assist (Ishida & Gayko, 2004; Young & Stanton, 2007) are introduced as standard features on many contemporary vehicles . Vehicle manufacturers are trying to combine these function specific assistance systems (NHTSA, 2013) into a holistic solution, called combined function assistance (NHTSA, 2013) or Highly Automated Driving (HAD). Examples of such technology emerging into the marketplace include 'Integrated Cruise Assist' (Bosch, 2015), 'Autopilot' (Tesla Motors, 2016), 'Intellisafe Autopilot' (Volvo Cars, 2016) and 'Highway Pilot' (Daimler, 2016). These systems automate both longitudinal and lateral aspects of driving, as well as automating some of the traditional decision-making tasks of the driver, such as anticipation of velocity reduction, monitoring lane position, and adherence to speed limitations (Banks et al., 2014; Kircher et al., 2014; Stanton et al., 1997; Stanton & Young, 2005). This is a form of "driver initiated automation", where the driver is in control of when the system is engaged or disengaged (Banks & Stanton, 2015, 2016; Lu & de Winter, 2015). Such HAD systems could enable the driver to become hands-free and feet-free (Banks & Stanton, 2014).

One of the main benefits of HAD is its potential for reducing the number of road traffic accidents. In 2010, NHTSA reported that the cost of motor vehicle crashes amounted to $242 billion per annum and 32,999 fatalities in the United States (Blincoe et al., 2015), and over 1.2 million fatalities worldwide (World Health Organization, 2009). Elon Musk, CEO of Tesla Motors stated that *"The probability of having an accident is 50% lower if you have Autopilot on. Even with our first version. So we can see basically what's the average number of kilometers to an accident – accident defined by airbag deployment. Even with this early version, it's almost twice as good as a person."- Musk (2016).* Furthermore, Ross (2016) showed that Tesla Autopilot maintains its distance to the lane centre more consistently than manual drivers. Whilst it remains to be seen whether HAD features can yield significant decreases in accident rates (Kalra & Paddock, 2016), it is estimated that HAD could greatly reduce societal costs such as medical, legal, emergency service (EMS), insurance administration and congestion costs, property damage, and workplace losses resulting from accident involvement (Blincoe et al., 2015). This could help progress towards the goal of the European Commission to halve the number of road deaths in the European Union by 2020 (European Commission, 2010).

Even so, HAD should not be viewed as a panacea in driving safety (Kalra & Paddock, 2016). HAD features are unable to cope with all possible driving scenarios. This was demonstrated by the recent Tesla incident where a vehicle crashed in to a trailer with the Autopilot engaged (Levin & Woolf, 2016). HAD features operate within strict functional limits and once these limits are reached, ceases to function effectively, if at all (SAE J3016, 2016; Stanton, 2015). Despite the good intentions of HAD, the sudden increase in demand resulting from a transition between HAD to manual control (De Winter et al., 2014; Stanton et al., 1997), could pose a significant problem for drivers of HAD vehicles as driving is a very demanding activity that comprises of over 1600 sub-tasks (Walker et al., 2015).

Human Factors research into automated driving has been ongoing since the mid-90s (Nilsson, 1995; Stanton & Marsden, 1996). As the motor-industry advances toward HAD, research conducted in driving simulators will become ever more important (Boer et al., 2015). Driving simulators have the advantage of allowing the evaluation of driver reactions to new technology within a virtual environment without the physical risk found on roads (Carsten & Jamson, 2011; De Winter et al., 2012; Flach et al., 2008; Stanton et al., 2001; Underwood et al., 2011). It is widely accepted that driving simulation offers a high degree of controllability and reproducibility as well as providing access to variables that are difficult to accurately determine in the real world (Godley et al., 2002), such as lane position and distance to roadway objects (Santos et al., 2005; Van Winsum et al., 2000).

When evaluating the validity of a simulator, Blaauw (1982) distinguished between two types of simulator validity; physical and behavioural validity. Physical validity refers to the level of correspondence between the physical layout, the configuration of the driver cabin, components and vehicle dynamics in the simulator and a real world counterpart. Behavioural fidelity, or the correspondence in driver behaviour between the simulator and its on-road counterpart, is arguably the most important form of validity when it comes to the evaluation of a specific task (Blaauw, 1982). Behavioural fidelity can be further extended into absolute validity and relative validity. Absolute validity is obtained when the absolute size of an effect measured in a simulator is the same as the absolute effect measured in its on-road counterpart. Relative validity on the other hand describes how well the relative size, or direction of an effect measured in the simulator corresponds to real driving (Blaauw, 1982; Kaptein et al., 1996).

There is plenty of research the design of Human Machine Interfaces, driver errors and task load, very little of the research has demonstrated transfer from the simulated environment to the open road (Mayhew et al., 2011; Santos et al., 2005; Shechtman et al., 2009; Stanton et al., 2011; Stanton & Salmon, 2009; Stanton et al., 2001; Wang et al., 2010). Most of the research showing how drivers interact with highly automated vehicles outside of simulators have taken place on closed test tracks (Albert et al., 2015; Llaneras et al., 2013; Stanton et al., 2011). Only a minority of studies on HAD being performed on the road (Banks & Stanton, 2016). Those remaining studies have investigated sub-systems such as Adaptive Cruise Control (Beggiato et al., 2015; Morando et al., 2016) and Lane Keeping Assistance systems (euroFOT, 2012; Ishida & Gayko, 2004; Stanton et al., 2001). This means that there is a paucity of research into the relative validity of driver behaviour in simulated HAD vehicles. This lack of studies could be attributed to the costs and risks associated with non-professional drivers driving prototype vehicles (such as the Mercedes S/E-class and Tesla vehicles equipped with these features, for road testing (Mercedes-Benz, 2015; safecarnews.com, 2015; Tesla Motors, 2016)) Consequentially, most research into human-automation interaction has been limited to simulators (for a review on control transitions in the simulator see Eriksson & Stanton, 2017) or closed test tracks (e.g. Albert et al., 2015; Llaneras et al., 2013; Stanton et al., 2011). A disadvantage of testing on closed test track compared to on road testing is the reduced complexity and dissonance between driver behaviour on the track and normal on road driving as well as the lack of other road users.

The purpose of the research reported in this paper was to explore whether control transitions between automated driving and manual driving observed in a driving simulator study are similar to real-world driving. A recent meta-analysis found that drivers of manual vehicles (SAE Level 0) take approximately 1 second to respond to sudden events in traffic (Eriksson & Stanton, 2016). It was also found that drivers of "function specific automation" (ACC and assistive steering, SAE Level 1 and 2) took an additional 1.1-1.5 seconds to respond to a sudden automation failure and that drivers of HAD vehicles (SAE level 3) took on average 2.96±1.96 seconds to respond to a control transition request leading up to a critical event, such as a stranded vehicle (Eriksson & Stanton, 2016). In contrast, Google (2015) reported that it takes their professional test drivers 0.84 seconds to respond to automation failures of their autonomous (SAE Level 4/5) prototypes whilst driving on public roads based on 272 discrete events. Moreover, the meta-analysis showed that the response time varies with the lead-time between the control transition request and a critical event. The reported lead times to the critical event at the point the request from manual control was issued varied between 2 and 30 seconds, and was 6.37 seconds on average. This is somewhat problematic as the SAE guidelines for level 3 automation states that the driver: *"Is receptive to a request to intervene and responds by performing dynamic driving task fallback in a timely manner" (SAE J3016, 2016, p. 20).* A decision to explore control transitions in non-urgent situations was made due to the lack of research into driver-paced transitions of control, which arguably is one of the more common use-cases for HAD control transitions, when for example leaving a highway.

2 Method

This paper is based upon the results of a two-phase between-participant research project. The first phase involved collecting times for control transitions within a simulated driving environment and the second phase collected the same data from the open road. The experimental design and procedure for each study are discussed in turn.

2.1 Phase 1

Participants

Phase one of the study used 26 participants (10 females, 16 males) between 20 and 52 years of age (Mean = 30.27 *SD* = 8.52) with a minimum one year driving experience (Mean = 10.57, *SD* = 8.61). This part of the study had been approved by the Southampton University ERGO ethics committee (RGO number 17771). Participants had no previous experience with ADAS systems.

Equipment

The study was carried out in a fixed based driving simulator located at the University of Southampton. The simulator was a full cab Jaguar XJ 350 with integrated pedal and steering sensors provided by Systems Technology Inc. as part of STISIM Drive® M500W Version 3 (http://www.stisimdrive.com/m500w) providing a projected 140° field of view. The instrument cluster was displayed on a 10.6” Sharp LQ106K1LA01B Laptop LCD panel display fitted in place of the original instrument cluster. Participants were instructed to drive at a speed of 70 mph on a 30 kilometre, three lane highway with some curves, with oncoming traffic in the opposing three lanes separated by a barrier and moderate traffic conditions.

2.2 Phase 2

Participants

The second phase of the study comprised of 12 participants (6 males, 6 females) between 20 and 49 years of age (Mean = 32.33 SD = 10.98) with a minimum one year driving experience (Mean = 14.58, SD = 11.13). All participants in the on road trial had undertaken extended driver training as a legal requirement for insurance purposes for the execution of phase 2; and therefore had previous experience with Advanced Driver Assistance Systems, such as Adaptive Cruise Control or Lane Keeping Assist. Nevertheless, none of the drivers had previous experience with the Tesla Autopilot system. The second phase of the study was approved by the Southampton University ERGO ethics committee (RGO number 19151).

Equipment

Phase 2 of the study was conducted using a Tesla Model S P90 equipped with the Autopilot software feature which enables short periods of hands- and feet-free driving on motorways as longitudinal and lateral control becomes automated. Drivers were reminded that they were ultimately responsible for safe vehicle operation and were not actively encouraged to remove their hands from the wheel at any point during the study. To ensure consistency between the two experiments, an iPad was mounted next to the instrument cluster running the application "Duet Display". This enables the iPad to act as a secondary monitor, displaying the same type of 'Take-over request' (TOR) visual feedback and auditory messages as in the simulator trial. The TOR's were reset by the experimenter, sat in the rear of the vehicle, in a Wizard-of-Oz fashion (Dahlbäck et al., 1993). To capture the control transitions a Video VBox Pro from Racelogic was used.

Participants were invited to drive along public roads and highways within Warwickshire, United Kingdom (B4100, M40 and M42). They were asked to adhere to national speed limits at all times and to keep lane changes to a minimum. Data recording of TOR response times took place on the M40 and M42 where speed was limited to 70 mph.

2.3 Procedure

Upon providing informed consent, participants in both phases of the study were provided with additional information about the HAD feature they would be driving with. They were told that the Tesla system could be overridden via the steering wheel, throttle or brake pedals, and through a touch screen interface in the simulator. Participants were reminded that they were responsible for the safe operation of the vehicle at all times, regardless of its mode (manual or automated) in accordance with recent amendments to the Vienna Convention of Road Traffic (United Nations, 1968). Participants were told that the system may prompt them to either resume or relinquish control of the vehicle during the drive, and that they should adhere to the instruction only when they felt it was safe to do so. This was intended to reduce the pressure on participants to respond immediately and to reinforce the idea that they were ultimately responsible for safe vehicle operation. For the 12 participants involved in the on-road study, additional instructions were given to ensure they remained aware of the vehicle's internal HMI (specifically about the state of the Autopilot) in an effort to maintain the safety of the vehicle driver and passengers in case of Autopilot malfunction, or failure to engage the Autopilot due to for example, missing lane markings. To support them in doing this, a qualified safety driver was present in the passenger seat at all times, ready to prompt the drivers to take back control if the need to regain control arose, or to press the emergency stop button in the Tesla centre display should the driver be unresponsive to prompts by the safety driver.

In both phases of the study, control transition requests were presented as both a visual cue (Figure 1) and an auditory message in the form of a computer-generated, female, voice stating *"please resume control"* or *"automation available".* The interval in which these requests were issued ranged from 30-45 seconds, thus allowing for approximately 24 control transitions half of which were to manual during the approximately 20 minute drive on the M40 and M42.

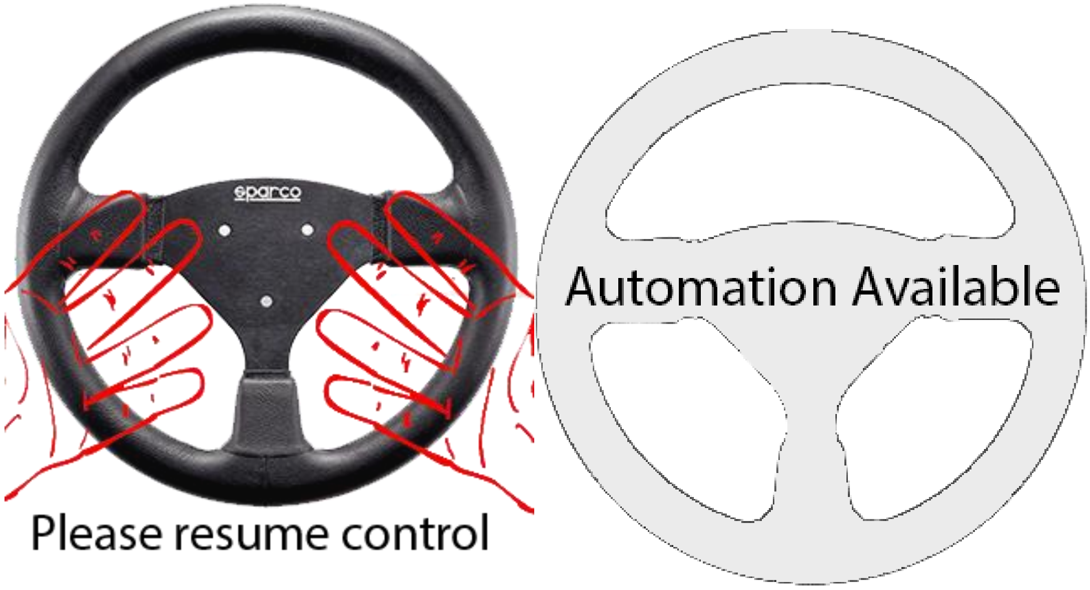

Figure 1. Left hand side, the instrument cluster showing a take-over request. The visual TOR was coupled with a computer-generated voice message stating "please resume control". On the right hand side is a control transitions request to automated vehicle control presented in the instrument cluster, coupled with a computer-generated voice message stating "automation available".

In the simulated driving condition the HMI used to switch mode was located in the centre display, running on a windows tablet, consisting of two buttons used to engage or disengage the automated driving feature. The automated driving system in the simulator was set to disengage only when the mode-switching buttons were pressed to allow for consistent control transitions. In the on-road driving condition the automated driving feature was engaged by a 'double pull' on a control stork on the left side of the steering wheel, below the indicator stork. To disengage the automated driving feature the driver could either; depress the brake, to disengage both the ACC and Lateral control, apply a steering input to disengage the lateral control only, or press the control stork forwards to disengage the ACC and lateral control.

Reaction times to the control transition request were recorded for each participant. In phase 1, reaction time was recorded from the onset of stimuli until the driver completed the requested action. In phase 2, reaction times to control transition requests were captured through Racelogic video VBOX Pro and manually coded based upon the mode-indicator in the Tesla Instrument Cluster switching mode after the control transition request was displayed in the iPad display. At the end of each drive, participants were asked to fill out the NASA-TLX (Hart & Staveland, 1988) and Technology acceptance Scale (Van Der Laan et al., 1997) with respect to the control transition process.

3 Analysis

The median Take Over Reaction Time values for each participant were calculated (Baayen & Milin, 2010) after which Wilcoxon rank sum tests were computed to analyse response-time and TLX data. The box plots in Figure 2 were adjusted to accommodate the log-normal distribution of the Take Over Response-time data (Hubert & Vandervieren, 2008).

To enable correlation analysis of take-over response times between the two groups of participants, drivers from the simulated drive were matched with drivers from the on road driving scenario on gender, age and driving experience as shown in Table 1. An uneven sample size was still present after participant matching, with fewer transitions recorded in the on road condition. Therefore, a randomised removal of observations on a participant-by-participant basis was conducted to ensure equal number of observations for each participant pair. After data reduction, the take-over response times for each task condition were added to a single vector and sorted in ascending order which, according to Ryan and Joiner (1976), enables a comparison of the two distributions to be made.

Table 1. Participants matched on age and gender from the on road and simulator experiments. The participants are ordered based on participant number for the on-road trial.

		On road		Simulator	
Participant	Gender	Age	Driving experience (Yrs)	Age	Driving experience (Yrs)
1	Male	20	3	20	1
2	Male	24	3	24	7
3	Female	29	12	29	6
4	Male	59	43	52	35
5	Female	32	15	28	10
6	Male	26	9	24	6
7	Male	28	11	28	6
8	Female	30	12	28	10
9	Female	32	14	40	22
10	Female	49	28	50	33
11	Female	26	8	27	10
12	Male	33	17	34	17
Mean		32.33	14.58	32	13.58
SD		10.98	11.13	10.20	10.99

4 Results

The results showed significant differences between on road, and simulated driving when relinquishing control to the vehicle automation ($Z = -6.120$, $p < 0.01$). On average, a one second increase in response time was found (see Table 2, Figure 2) when relinquishing control to the vehicle automation in the simulated road condition. It took drivers approximately 3.18±2.83 seconds to relinquish control in the on-road condition, whilst it took 4.20±1.96 seconds to relinquish control in the simulated driving condition.

Table 2. Control transition times (seconds) between manual and automated and automated and manual control in the on road, and simulator condition

	To Automated		To Manual	
	Simulator	On road	Simulator	On road
Median	4.20	3.18	4.56	3.08
IQR	1.96	2.83	1.63	1.16
Min	2.82	1.33	1.97	1.21
Max	23.88	25.16	25.75	15.41

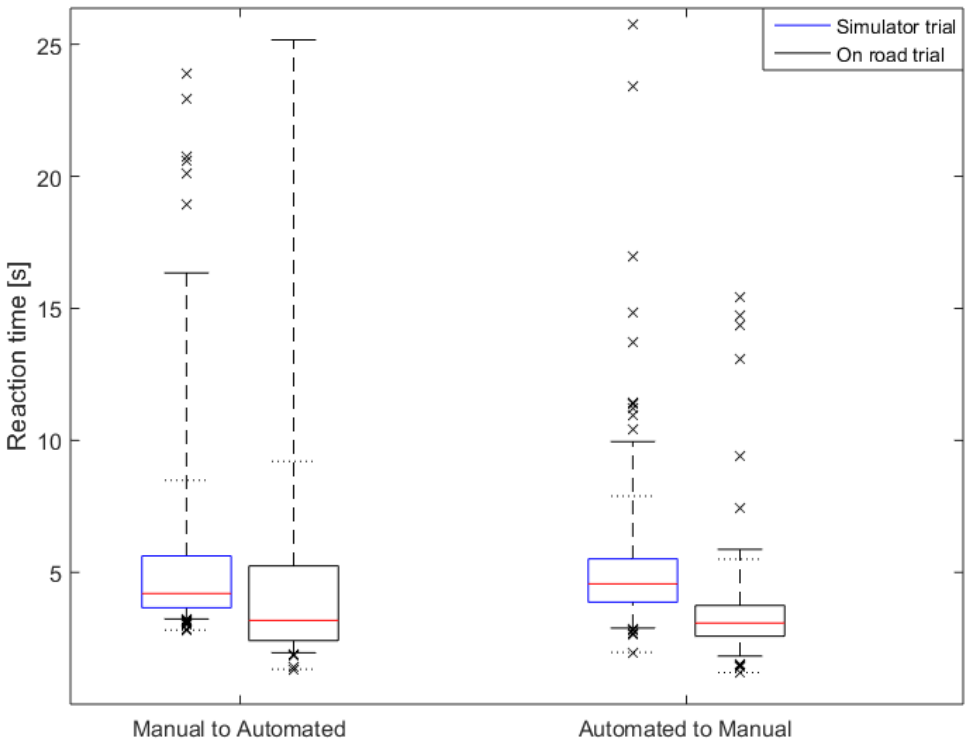

Figure 2. Control transition times from manual to automated control and from automated to manual control in the on road and simulated driving condition

Moreover, a significant decrease in take-over request reaction time of approximately 1.5 seconds was found in the on road driving condition, in comparison to the simulated driving condition when resuming control from the automation (Z = -10.403, $p < 0.01$). It took drivers approximately 3.08±1.16 seconds to resume control from the automation in the on road driving condition and 4.56±1.63 seconds to resume control in the simulated driving condition.

The results from the correlation analysis of the sorted response time data for the transition to automated driving from manual driving showed a strong positive correlation (Pearson's r=0.96, p<0.0001, calculated power = 1.0), as illustrated in Figure 4.

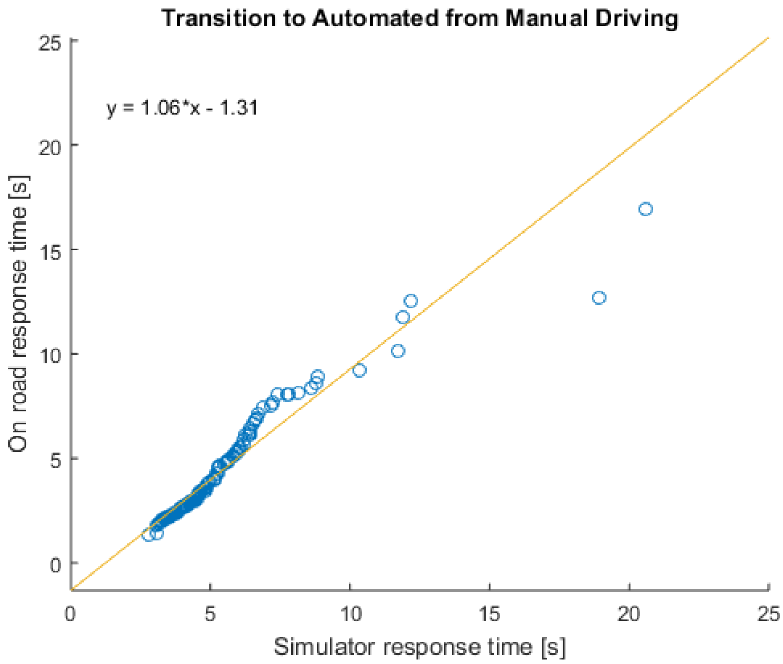

Figure 3. Scatter plot of control transitions from manual to automated vehicle control. The X axis shows the driver response time in the simulator, the Y axis show the driver response time in the on-road condition.

The correlation analysis of the control transition time from automated to manual control showed a significant positive relationship (Pearson's r=0.97, p<0.0001, calculated power = 1.0) between the two distributions as shown in Figure 5.

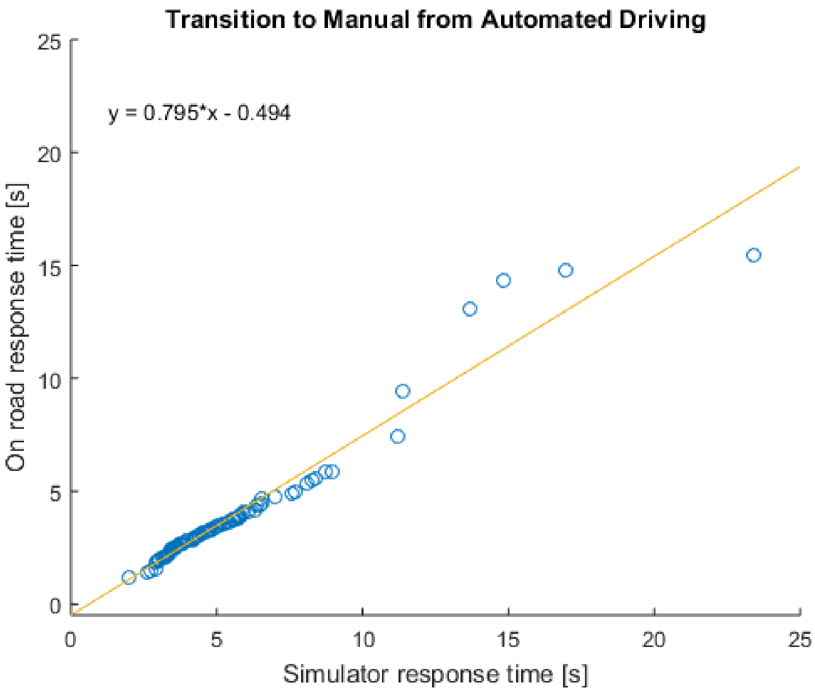

Figure 4. Scatter plot of control transitions from automated to manual vehicle control. The X axis shows the driver response time in the simulator, the Y axis show the driver response time in the on-road condition.

Subjective workload scores collected through the NASA-TLX sub-scales (Byers et al., 1989; Hart & Staveland, 1988) at the end of each driving condition showed no significant correlations on the workload sub-scales nor overall workload for the matched sample. A comparison of the two conditions showed no significant differences on any of the sub-scales or overall workload (Table 3). Overall there was little difference in workload, and the median workload in both conditions was approaching the halfway point on the scale (Figure 8), implying relatively low workload.

Table 3. NASA-TLX and Technology acceptance scores scores of the on road and simulated driving conditions. The effect size was calculated as r=abs(Z)/√(N1+N2)

	On road	Simulator	Rank sum test			Pearson's Correlation	
	Median (SD)	Median (SD)	Z	P	r	p	r
Mental Demand	8 (7)	7.5 (10)	-0.603	0.53	0.09	0.60	-0.16
Physical Demand	2.5 (4.5)	4(5)	1.395	0.16	0.23	0.71	-0.12
Temporal Demand	4.5(5)	3(6)	-0.415	0.68	0.06	0.81	-0.08
Performance	5(2.5)	6(7)	0.872	0.38	0.14	0.81	-0.08
Effort	5.5(5)	5(9)	0.206	0.84	0.03	0.66	0.14
Frustration	4.5(4)	4(9)	0.269	0.79	0.04	0.25	-0.35
Overall workload	4.75(4.2)	5.16 (7.33)	0.755	0.45	0.12	0.92	-0.03
Usefulness	1.1 (0.6)	1 (1)	0.491	0.62	0.07	0.64	-0.14
Satisfaction	1 (0.87)	0.5(1.69)	1.316	0.18	0.21	0.43	-0.25

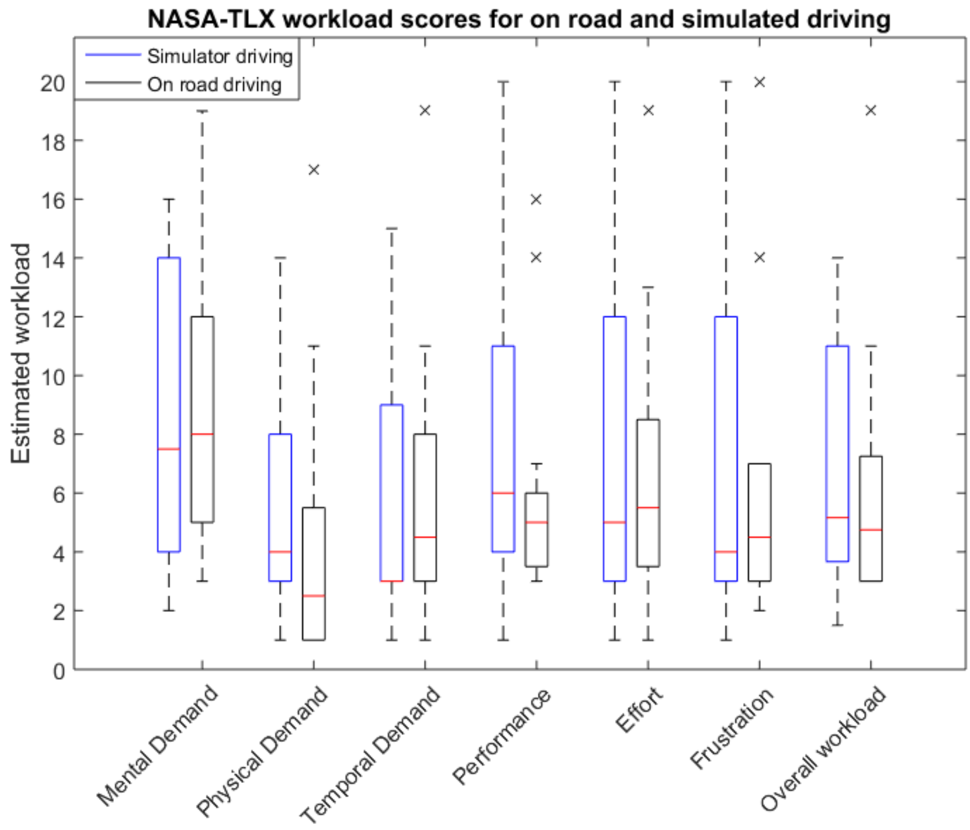

Figure 5. Self-reported workload scores for the simulated drive and the on road drive

The Van Der Laan Technology Acceptance Scale yielded no significant correlations on the matched samples as shown in Table 3. Moreover, no significant differences in automation usefulness (Z= 0.491

$p > 0.05$, $r = 0.07$) and automation satisfaction ($Z = 1.316$ $p > 0.05$, $r = 0.21$) between the on road and simulated driving conditions could be found.

5 Discussion

As shown by Eriksson and Stanton (2017) most research into control transitions in Highly Automated Driving has been undertaken in driving simulators. The studies outside of simulators tend to be limited to closed test tracks (Albert et al., 2015; Llaneras et al., 2013), or to sub-systems of HAD (Beggiato et al., 2015; euroFOT, 2012; Ishida & Gayko, 2004; Morando et al., 2016). One study by Banks and Stanton (2016) has explored the interaction with automated vehicles on the open road, but links to performance on similar tasks in simulated environments were not made. To further the understanding of the validity of driving simulators in Highly Automated Driving we compared multiple control transitions; from manual driving to automated driving and vice versa, for both simulated, and on-road driving environments. The reason for the unusually high frequency of transitions of control compared to contemporary literature was to reduce the impact of novelty effects, to enable the capturing of inter- and intra-individual differences to get an appreciation for the wide range of transition times, and to compress experience with the system as previously done by Stanton et al. (2001). Moreover, it was argued in Eriksson and Stanton (2017) that the type of control transition utilised in this manuscript is 'non-urgent' and such transitions are likely to be commonplace on public roads when SAE level 3 systems are limited to certain operational constraints (SAE J3016, 2016).

The results show that drivers in the on road driving condition took on average 3.08 seconds. This is marginally longer compared to the average 2.96 control resumption time for drivers who are required to resume control within a limited time-frame (e.g. 7 seconds as in Gold et al. (2013)). It has previously been shown that permitting drivers to self-regulate the use of in-vehicle technologies on a tactical level tend to maintain optimal workload and safer driving performance (Cooper et al., 2009; Eriksson et al., 2014; Eriksson & Stanton, 2017; Kircher et al., 2016; Young & Stanton, 2007). Moreover, the results show that there is a long tail in the resumption-time distribution indicating that some drivers take up to 15 seconds to resume control. This shows that the design of the take-over process should not focus on average resumption times, but rather use the 5^{th} – 95^{th} percentile user, as is common practice in anthropometrics (Eriksson & Stanton, 2017; Porter et al., 2004). The results also show that drivers were generally faster in the execution of control transitions in the on-road driving scenario for transitions to both automated, and manual, control compared to the simulated driving scenario. Similar effects have been observed by Wang et al. (2010) and Kurokawa and Wierwille (1990) where drivers produced faster responses in on-road conditions than in the simulator for in-vehicle interaction tasks.

This difference could be partially explained by the perception of greater risk in the on-road condition (Carsten & Jamson, 2011; De Winter et al., 2012; Flach et al., 2008; Underwood et al., 2011). Moreover, the differences between mode switching human-machine interface in the Tesla (control stork next to the steering wheel) and the simulator (touch screen in centre console) could account for the increased response time in the simulator part of this study. Drivers have been found to have significantly higher eyes-off-road time when engaged with in-vehicle systems with high visual demands, and as driving is a visually demanding task this can have large effects on driving performance (Jæger et al., 2008). This increase in eyes-off-road could be further amplified by the virtue of using a touch screen which lacks the haptic nature of standard vehicle interface elements, such as the control stork in the Tesla, that enables blind interaction whilst driving (Rümelin & Butz, 2013). It could be that drivers had to divert visual resources to identify which of the two buttons to press to reach the desired state, and to plan a motor-path to execute the action to press the button. For experienced drivers this is a well-practiced behaviour, executed whenever a driver needs to change the radio station, confirm a rerouting on their sat-nav, or change the heating settings of their vehicle. It can therefore be argued that this type of interaction should have a negligible effect on the transition times compared to the magnitude of effects observed in the literature of driver reaction time on the road, and in the simulator (Kurokawa & Wierwille, 1990; Wang et al., 2010).

Another factor that could have influenced take-over request response time could be the different levels of experience with Advanced Driver Assistance Systems between the two samples, where drivers in the experienced (Tesla) group produced faster reaction times due to their familiarity with such systems. It is important to acknowledge these factors as they may have contributed to the faster response times in the on road condition. However, as effects of a similar magnitude have been found previously (Kurokawa & Wierwille, 1990; Wang et al., 2010) it stands to reason that the main part of the observed difference can be accounted for by the nature of the driving environment, and to a lesser extent, the mode switching HMI and the differences in ADAS experience. The consistent difference in response times between the two conditions, suggest that absolute validity could not be found. Nevertheless, evidence for relative validity between the on-road and simulated environments is strong as the correlation analysis showed a strong positive similarity of the distributions for both types of control transitions. The lack of correlations in workload and technology acceptance scores could be explained by the subjective nature of the questionnaire, combined with a relatively small sample in a between-group study-design. It is worth noting that the workload and technology acceptance scores are not absolute scores, and therefore need to be looked upon with some caution due to the relatively small sample of two independent groups.

Despite the lack of correlation between the on-road and simulated drive on the subjective questionnaires, the lack of differences in workload and Technology acceptance scores, indicate that the driving conditions had no measurable effect on perceived workload, usefulness and satisfaction. We therefore argue that relative fidelity of the simulator can be established with regard to human-automation interaction and, in particular, control transitions (Blaauw, 1982; Godley et al., 2002). These results support Stanton et al. (2001) who found that driver performance on secondary tasks were highly correlated when performed on the road and in the simulator during manual driving.

Consequentially, the results obtained in this study lends validity to previous research into control transitions in automated vehicles carried out in simulated environments. In light of these results, researchers may have more confidence when using simulators as a primary tool for research on human-automation interaction (Stanton et al., 2001). This observation permits the exploration of phenomena related to automated vehicles in a reproducible, deterministic, and completely observable environment (Russel & Norvig, 2009), and facilitates the collection of data that would otherwise be difficult to obtain in road vehicles (Godley et al., 2002; Santos et al., 2005; Van Winsum et al., 2000). These findings show that driving simulators are legitimate tools for researching vehicle automation (Boer et al., 2015).

6 Conclusions

In this paper the validity of human-automation interaction in highly automated vehicles in driving simulators was assessed. Absolute validity could not be established due to the shorter transition times observed in the on road driving condition. It was found that on average drivers take an additional second to transfer control to the automation in the simulated drive, and an additional 1.4 seconds to resume control in the simulated drive compared to on-road. Moreover, it was found that drivers in the on-road driving condition were marginally faster than what has been found in previous literature when drivers have resumed control under time-pressure. Despite these similarities, it was also shown that there is a long tail in the distribution of resumption-times, and that these drivers will have to be accommodated to ensure safe use of automation. Nevertheless, the results also showed that there was a strong positive correlation for transition time in the on-road and simulated driving conditions. In light of these results the authors argue that there is a strong indication of relative validity for research conducted in simulators. Despite the lack of significant correlations we argue that relative validity is further supported by the similarities in workload, and technology acceptance scores of the drivers in the simulated, and on road driving conditions.

Consequentially, in this study, the authors argue that the driving simulator is a valid research tool for the exploration human-automation interaction, and in particular the transfer of control between driver and automation. In conclusion, medium-fidelity, fixed based, driving simulation is a safe and cost-effective method for assessing human-automation interaction, and in particular control transitions in highly automated driving.

Acknowledgements

This research has been conducted as a part of the European Marie Curie ITN project HFAuto - Human Factors of Automated driving (PITN-GA-2013-605817). The authors would like to thank Jim O'Donoghue and Jaguar Land Rover Ltd. for making this study possible.

References

Albert, M., Lange, A., Schmidt, A., Wimmer, M., & Bengler, K. (2015). Automated driving–Assessment of interaction concepts under real driving conditions. *Procedia Manufacturing,* vol *3*, pp 2832-2839.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research,* vol *3*, no 2, pp 12-28.

Banks, V. A., & Stanton, N. A. (2014). Hands and Feet Free Driving: Ready or Not? In N. S. S. Landry, G. D. Bucchianico & A. Vallicelli (Eds.), *Advances in Human Aspects of Transportation: Part II*: AHFE Conference.

Banks, V. A., & Stanton, N. A. (2015). Contrasting models of driver behaviour in emergencies using retrospective verbalisations and network analysis. *Ergonomics,* vol *58*, no 8, pp 1337-1346.

Banks, V. A., & Stanton, N. A. (2016). Keep the driver in control: Automating automobiles of the future. *Appl Ergon,* vol *53 Pt B*, pp 389-395.

Banks, V. A., & Stanton, N. A. (2017). The Unknown Paradox of "Stop the Crash" Systems: Are We Really Improving Driver Safety? In N. A. Stanton, S. Landry, G. D. Bucchianico & A. Vallicelli (Eds.), *Advances in Human Aspects of Transportation* (pp. 525-533): Springer.

Banks, V. A., Stanton, N. A., & Harvey, C. (2014). Sub-systems on the road to vehicle automation: Hands and feet free but not 'mind' free driving. *Safety Science,* vol *62*, pp 505-514.

Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour,* vol *35*, pp 75-84.

Blaauw, G. J. (1982). Driving experience and task demands in simulator and instrumented car: a validation study. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol *24*, no 4, pp 473-486.

Blincoe, L., Miller, T., Zaloshnja, E., & Lawrence, B. (2015). The economic and societal impact of motor vehicle crashes 2010. (Revised) (Report No DOT HS 812013). *Washington, DC: National Highway Traffic Safety Administration*.

Boer, E. R., Della Penna, M., Utz., H., Pedersen, L., & Sierhuis, M. (2015). *The role of driving simulators in developing and evaluating autonomous vehicles.* Paper presented at the the Driving Simulation Conference Europe Paris, France.

Bosch. (2015). *Making the most of car time*. http://annual-report.bosch.com/AUTOMATED-DRIVING.

Byers, J. C., Bittner, A., & Hill, S. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary. *Advances in industrial ergonomics and safety I*, pp 481-485.

Carsten, O., & Jamson, A. (2011). Driving simulators as research tools in traffic psychology. In B. Porter (Ed.), *Handbook of Traffic Psychology* (Vol. 1, pp. 87-96): Academic Press.

Cooper, J. M., Vladisavljevic, I., Medeiros-Ward, N., Martin, P. T., & Strayer, D. L. (2009). An investigation of driver distraction near the tipping point of traffic flow stability. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol *51*, no 2, pp 261-268.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). *Wizard of Oz studies: why and how.* Paper presented at the Proceedings of the 1st International conference on Intelligent User Interfaces, Orlando, Florida, USA — January 04 - 07, 1993

Daimler. (2016). The auto pilot for trucks: Highway Pilot. *https://www.daimler.com/innovation/autonomous-driving/special/technology-trucks.html*.

De Winter, J. C., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation research part F: traffic psychology and behaviour,* vol *27*, pp 196-217.

De Winter, J. C. F., van Leeuwen, P., & Happee, R. (2012). *Advantages and Disadvantages of Driving Simulators: A Discussion.* Paper presented at the Measuring Behavior, Utrecht, The Netherlands, August 28-31, 2012.

Eriksson, A., Lindström, A., Seward, A., Seward, A., & Kircher, K. (2014). Can user-paced, menu-free spoken language interfaces improve dual task handling while driving? *Human-Computer Interaction. Advanced Interaction Modalities and Techniques* (pp. 394-405): Springer.

Eriksson, A., & Stanton, N. A. (2017). Take-over time in highly automated vehicles: non-critical transitions to and from manual control. *Human Factors*.

euroFOT. (2012). euroFOT – Bringing intelligent vehicles to the road. *Retrieved from: http://www.eurofot-ip.eu/*.

European Commission. (2010). *Towards a European road safety area: policy orientations on road safety 2011-2020*: European Commission.

Flach, J., Dekker, S., & Stappers, P. J. (2008). Playing twenty questions with nature (the surprise version): reflections on the dynamics of experience. *Theoretical Issues in Ergonomics Science,* vol *9*, no 2, pp 125-154.

Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accid Anal Prev,* vol *34*, no 5, pp 589-600.

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). *"Take over!" How long does it take to get the driver back into the loop?* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Google. (2015). *Google Self-Driving Car Testing Report on Disengagements of Autonomous Mode*. https://static.googleusercontent.com/media/www.google.com/en//selfdrivingcar/files/reports/report-annual-15.pdf accessed on 01/06/2016.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology,* vol *52*, pp 139-183.

Hubert, M., & Vandervieren, E. (2008). An adjusted boxplot for skewed distributions. *Computational statistics & data analysis,* vol *52*, no 12, pp 5186-5201.

Ishida, S., & Gayko, J. E. (2004). *Development, evaluation and introduction of a lane keeping assistance system.* Paper presented at the Intelligent Vehicles Symposium, 2004 IEEE.

Jæger, M. G., Skov, M. B., & Thomassen, N. G. (2008). *You can touch, but you can't look: interacting with in-vehicle systems.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Kalra, N., & Paddock, S. M. (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice,* vol *94*, pp 182-193.

Kaptein, N., Theeuwes, J., & Van Der Horst, R. (1996). Driving simulator validity: Some considerations. *Transportation Research Record: Journal of the Transportation Research Board*, no 1550, pp 30-36.

Kircher, K., Eriksson, O., Forsman, Å., Vadeby, A., & Ahlstrom, C. (2016). Design and analysis of semi-controlled studies. *Transportation research part F: traffic psychology and behaviour*.

Kircher, K., Larsson, A., & Hultgren, J. A. (2014). Tactical driving behavior with different levels of automation. *IEEE Transactions on Intelligent Transportation Systems,* vol *15*, pp 158-167.

Kurokawa, K., & Wierwille, W. W. (1990). *Validation of a driving simulation facility for instrument panel task performance.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, October 1990 Santa Monica, CA, SAGE Publications.

Larsson, A. F. L., Kircher, K., & Andersson Hultgren, J. (2014). Learning from experience: Familiarity with ACC and responding to a cut-in situation in automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour,* vol *27, Part B*, pp 229-237.

Levin, S., & Woolf, N. (2016). Tesla driver killed while using autopilot was watching Harry Potter, witness says.   Retrieved 24 Aug 2016, from https://www.theguardian.com/technology/2016/jul/01/tesla-driver-killed-autopilot-self-driving-car-harry-potter

Llaneras, R. E., Salinger, J., & Green, C. A. (2013). *Human factors issues associated with limited ability autonomous driving systems: Drivers' allocation of visual attention to the forward roadway.*

Paper presented at the Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 92-98, Bolton Landing, New York,.

Lu, Z., & de Winter, J. C. F. (2015). A Review and Framework of Control Authority Transitions in Automated Driving. *Procedia Manufacturing,* vol *3*, pp 2510-2517.

Mayhew, D. R., Simpson, H. M., Wood, K. M., Lonero, L., Clinton, K. M., & Johnson, A. G. (2011). On-road and simulated driving: concurrent and discriminant validation. *J Safety Res,* vol *42*, no 4, pp 267-275.

Mercedes-Benz. (2015). Intelligent Drive next Level as part of Driving Assistance package. *https://www.mercedes-benz.com/en/mercedes-benz/innovation/with-intelligent-drive-more-comfort-in-road-traffic/*.

Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electron,* vol *38*, pp 114-117.

Morando, A., Victor, T., & Dozza, M. (2016). Drivers anticipate lead-vehicle conflicts during automated longitudinal control: Sensory cues capture driver attention and promote appropriate and timely responses. *Accident; Analysis and Prevention,* vol *97*, pp 206-219.

Musk, E. (2016). Interview at 'Future Transport Solutions' conference. *https://youtube/HaJAF4tQVbA?t=1341 Quote at 22 minutes 36 seconds: retrieved on 12 May 2016*.

National Highway Traffic Safety Administration. (2013). *Preliminary Statement of Policy Concerning Automated Vehicles*.

Nilsson, L. (1995). *Safety effects of adaptive cruise control in critical traffic situations.* Paper presented at the the Second World Congress on Intelligent Transport Systems: `Steps Forward`, Yokohama.

Porter, J. M., Case, K., Marshall, R., Gyi, D., & Sims neé Oliver, R. (2004). 'Beyond Jack and Jill': designing for individuals using HADRIAN. *International Journal of Industrial Ergonomics,* vol *33*, no 3, pp 249-264.

Ross, P. E. (2016). Tesla Reveals Its Crowdsourced Autopilot Data. *IEEE Spectrum. http://spectrum.ieee.org/cars-that-think/transportation/self-driving/tesla-reveals-its-crowdsourced-autopilot-data Accessed on 1/06/2016*.

Rümelin, S., & Butz, A. (2013). *How to make large touch screens usable while driving.* Paper presented at the Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Russel, S., & Norvig, P. (2009). *Artificial Intelligence A Modern Approach* (3rd edition ed.).

Ryan, T., & Joiner, B. (1976). Normal probability plots and tests for normality. Minitab Statistical Software: Technical Reports. The Pennsylvania State University, State College, PA. Available from MINITAB: Inc.

SAE J3016. (2016). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, *J3016_201609*: SAE International.

safecarnews.com. (2015). Intelligent Drive Concept for new Mercedes-Benz GLC. from http://safecarnews.com/intelligent-drive-concept-for-new-mercedes-benz-glc_ju6145/

Santos, J., Merat, N., Mouta, S., Brookhuis, K., & De Waard, D. (2005). The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. . *Transportation Research Part F* no 8, pp 135-146.

Seppelt, B. D., & Lee, J. D. (2007). Making adaptive cruise control (ACC) limits visible. *International Journal of Human-Computer Studies,* vol *65*, no 3, pp 192-205.

Shechtman, O., Classen, S., Awadzi, K., & Mann, W. (2009). Comparison of driving errors between on-the-road and simulated driving assessment: a validation study. *Traffic injury prevention,* vol *10*, no 4, pp 379-385.

Stanton, N. A. (2015, March). Responses to Autonomous Vehicles. *Ingenia,* no 62, pp 69.

Stanton, N. A., Dunoyer, A., & Leatherland, A. (2011). Detection of new in-path targets by drivers using Stop & Go Adaptive Cruise Control. *Appl Ergon,* vol *42*, no 4, pp 592-601.

Stanton, N. A., & Marsden, P. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science,* vol *24*, no 1, pp 35-49.

Stanton, N. A., & Salmon, P. M. (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. *Safety Science,* vol *47*, no 2, pp 227-237.

Stanton, N. A., Young, M., & McCaulder, B. (1997). Drive-by-wire: The case of driver workload and reclaiming control with adaptive cruise control. *Safety Science,* vol *27*, no 2-3, pp 149-159.

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics,* vol *48*, no 10, pp 1294-1313.

Stanton, N. A., Young, M. S., Walker, G. H., Turner, H., & Randle, S. (2001). Automating the Driver's Control Tasks. *International Journal of Cognitive Ergonomics*.

Tesla Motors. (2016). Model S Software Version 7.0. *vol* no, pp. Retrieved from https://www.teslamotors.com/presskit/autopilot

Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour,* vol *14*, no 6, pp 435-446.

United Nations. (1968). *Convention on road traffic*, Done at Vienna on 8 November 1968. Amendment 1. Retrieved from http://www.unece.org/fileadmin/DAM/trans/conventn/crt1968e.pdf.

Van Der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies,* vol *5*, no 1, pp 1-10.

Van Winsum, W., Brookhuis, K. A., & De Waard, D. (2000). A comparison of different ways to approximate time-to-line crossing (TLC) during car driving *Accident Analysis and Prevention, vol 32 (1)*, no 32, pp 47-56.

Volvo Cars. (2016). Intellisafe Autopilot. *http://www.volvocars.com/intl/about/our-innovation-brands/intellisafe/intellisafe-autopilot*.

Walker, G. H., Stanton, N. A., & Salmon, P. M. (2015). *Human Factors in Automotive Engineering and Technology*: Ashgate Publishing, Ltd.

Wang, Y., Mehler, B., Reimer, B., Lammers, V., D'Ambrosio, L. A., & Coughlin, J. F. (2010). The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing. *Ergonomics,* vol *53*, no 3, pp 404-420.

World Health Organization. (2009). *Global status report on road safety: time for action*: World Health Organization.

Young, M. S., & Stanton, N. A. (2007). Back to the future: brake reaction times for manual and automated vehicles. *Ergonomics,* vol *50*, no 1, pp 46-58.