A Longitudinal Simulator Study to Explore Drivers' Behaviour in Level 3 Automated Vehicles

David R. Large, Gary Burnett, Davide Salanitri
Human Factors Research Group,
University of Nottingham,
Nottingham, UK
{david.r.large; gary.burnett}@nottingham.ac.uk,
davide.salanitri@outlook.com

Anneka Lawson, Elizabeth Box
RAC Foundation,
London, UK
{Anneka.Lawson; Elizabeth.Box}
@RACFoundation.org

ABSTRACT
In a longitudinal study, 49 drivers undertook a commute-style journey, with part of the route supporting level-3 automation, over five consecutive days. Bespoke HMIs were provided to keep drivers in-the-loop during automation, and help them regain situational-awareness (SA) during handovers, in a 2×2 between-subjects design. Drivers demonstrated high levels of trust from the outset, delegating control to the vehicle (when available) and directing attention to their own activities/devices. Ratings of trust and technology acceptance increased during the week – even following an unexpected, emergency handover on day four – with the highest ratings recorded on day five. High levels of lateral instability were observed immediately following take-overs, although improvements were noted during the week and following the provision of SA-enhancing hand-over advice. Results demonstrate benefits associated with novel HMI designs to keep drivers in-the-loop and improve take-over performance, as well as the necessity of multiple exposures during the evaluation of future, immersive technologies.

Author Keywords
level 3 automation, trust, acceptance, driving performance, transfer of control, behavioural adaptation, HMI design

CSS Concepts
• Human-centered computing~User studies

INTRODUCTION
Autonomous or fully-automated vehicles (FAV) are expected to offer a number of benefits, including improvements in road safety, increased mobility, enhanced driver comfort and reductions in road congestion [1]. Relinquishing responsibility for vehicle control also allows drivers to use journey time for non-driving-related tasks, potentially providing a more enjoyable and productive experience during everyday car travel [2].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

AutomotiveUI '19, September 21–25, 2019, Utrecht, Netherlands
© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-6884-1/19/09…$15.00
https://doi.org/10.1145/3342197.3344519

However, FAV are not likely to populate our roads for quite some time [3], due to ongoing technical and legal complexities. In the meantime, there is an expectation that vehicles offering lower levels of automated control, or those that pose the ability to operate autonomously in certain situations only (e.g. 'geo-fenced' or traffic-jam assist-type technologies), are more likely.

The Society of Automotive Engineers (SAE) [4] categorises six levels of ascending automation (from level 0 to 5), that differ in the extent to which the system intervenes in vehicle control, and if and by how much the human driver needs to monitor the system (in anticipation of potentially taking over control). Intermediate, level 3 vehicles (also referred to as, 'conditional automation') are expected to be introduced onto UK public roads in the next few years [3]. However, at level 3, the human driver is still expected to be responsible for the vehicle's actions, and consequently must be available and prepared to resume manual control in situations the vehicle cannot handle. It has therefore been suggested that SAE level 3 automated vehicles allow drivers to "become hands and feet free, but not necessarily 'mind free'" [5]. Moreover, the situations in which drivers must regain manual control of automated vehicles may occur unexpectedly and require fast responses, and thus human drivers may be ill-prepared for the transfer of control [1, 5, 6].

A further concern is that drivers are likely to become 'out-of-the-loop' when they are not actively monitoring, making decisions or physical inputs to the driving task (i.e. not manually driving, themselves). This reduces their perception and comprehension of environmental elements and events, and their ability to project the future status of these items (i.e. their so-called 'situational awareness'). The driving skills hierarchy [7] identifies situational awareness as a key element at the 'tactical' level of driving. The hierarchy describes the relationship between strategic (planning), tactical (manoeuvring) and operational (control) elements of the driving task in a top-down relationship. The highest, 'strategic' level defines the overall journey goals and general plans, including route, mode choice etc. At the 'tactical' level, drivers negotiate the directly prevailing circumstances in controlled action patterns, including obstacle avoidance, gap acceptance, turning and overtaking etc., for which situational awareness is important. The ‘control’ level (at the bottom of the hierarchy) defines the physical control actions (or automatic action patterns) associated with safe vehicle manoeuvring (e.g. steering, braking, mirror checks etc.). To date, proposed take-over requests typically exist as a ‘bottom-up’ approach (with respect to Michon’s driving skills hierarchy [7]), demanding that the driver ‘take control’ (lower level skills), without attempting to assess or rebuild their situational awareness (higher level skills) prior to handing-over control.

A further concern is that, given the absence of responsibility for primary control actions during automated driving, human drivers will also likely engage in non-driving related activities, which could in and of themselves contribute to the loss of awareness of the vehicle system state and external driving environment (as drivers become cognitively captivated by their chosen activities). Previous, preliminary work has shown that given complete freedom to select and engage in activities of their choosing while the vehicle is in control – even when told that they may have to resume manual driving – drivers still chose highly-engaging activities, often with strong visual, manual and cognitive elements, and quickly become absorbed in these activities, at the apparent expense of their awareness of the external driving situation [2].

In addition, drivers engaged in non-driving related tasks during automated driving have been found to show greater signs of fatigue and mind wandering, and have slower reaction times to takeover requests, measured by the time taken to look at the road, place their hands on the steering wheel and their feet on the pedals (typical indicators of ‘readiness to drive’) [8].

Whilst these results make interesting reading, and offer valuable contributions in and of themselves, a potential criticism is that they were obtained from limited exposure – typically, a single-visit simulator study with less than an hour of total use (e.g. [1]). While there remains merit in this approach, the concern in this context is that such studies are unlikely to reveal behavioural adaptations to technologies, which would occur over time and with multiple exposures. This will be particularly relevant for future, self-driving vehicles, where the driver’s new role (which is yet to be fully established) and their overall experience will be fundamentally different to that in a manually-driven car, and drivers will therefore require time to adapt [6]. Thus, understanding how drivers’ behaviour changes over time could be pivotal to the effective design and ultimate acceptance of future vehicles.

Needless to say, it is inherently difficult to conduct empirical studies involving future technologies for which robust and reliable solutions are not yet widely available. In addition, there are obvious safety, ethical and legal restrictions associated with studies conducted on-the-road (particularly apropos of automated vehicles). Thus, the driving simulator presents itself as an ideal research tool in this context. Moreover, by utilising a longitudinal approach, involving multiple visits, it is possible to explore how attitudes and behaviour may change over time. The motivation behind the current study was therefore to extend the preliminary investigation conducted by Large et al. [2], by recruiting a larger cohort of drivers, and provide the opportunity (through repeated exposures) to explore behavioural adaptations to SAE level 3 driving over a week of use.

A further aim of the study was to explore possible Human-Machine Interface (HMI) interventions to help keep drivers ‘in-the-loop’ during periods of automation and to rebuild their situational awareness prior to the hand-over of control. For example, an HMI providing details of the current vehicle status (sensor operability etc.) could provide drivers with an indication of the vehicle’s ability to maintain control (potentially enabling drivers to predict an impending transfer of control), and a re-imagined ‘top-down’ take-over request (with respect to Michon’s [7] driving skills hierarchy) could provide ‘tactical’ information to guide the driver’s attention (aiming to increase their situational awareness), prior to the issuance of a take-over request at the ‘control’ level.

As with the preliminary investigation conducted by Large et al. [2], no restrictions were applied to the nature of activities with which participants could engage while the vehicle was in control – they were simply told to imagine what they might do in such a situation and bring any items they would require to accomplish this (mobile phone, laptop, book etc.): Naturally, participants were made aware that they may be required to resume manual control, given appropriate notice, in line with our understanding of the SAE definition of level 3 automation [4].

METHOD

Participants

Fifty-two participants were recruited to take part in the study via advertisements placed around the University of Nottingham campus and sent via email. All participants were active and competent drivers (more than two years driving, and drove regularly), and primarily comprised employees (administrative, academic, technical etc.) and postgraduate students at the University of Nottingham. Unfortunately, three participants were unable to complete the full study due to simulator sickness occurring partway through the week, leaving a total of 49 participants (27 male, 22 female; mean age: 32, range: 21-64; annual mileage: 5621). Participants were matched as closely as practicable between the four different conditions for age, gender and driving experience. They were reimbursed with £50 in shopping vouchers as compensation for taking part

Apparatus

The study took place in a medium-fidelity, fixed-base driving simulator at the University of Nottingham (Figure 1). The simulator comprises a right-hand drive Audi TT car positioned within a curved screen, affording a 270-degrees

Figure 1. University of Nottingham driving simulator, showing external (top) and internal (bottom) views

forward and contiguous side image of the driving scene via three overhead HD projectors, together with rear and side mirror displays. A Thrustmaster 500RS force feedback steering wheel and pedal set are integrated faithfully with the existing Audi primary controls, with the dashboard created using a bespoke application and re-presented on a 7-inch LCD screen, replacing the original Audi instrument cluster. During the study, the simulator was modified to mimic the capabilities of a SAE level 3-automated car, in so far as the driver was able to relinquish the physical primary control actions (steering, accelerating and braking) under certain conditions.

The simulated driving environment was created using STISIM (version 3) software (https://stisimdrive.com/), and was designed to replicate a typical 'commute' journey, lasting approximately 30 minutes. The route began in a residential location (described to participants as their 'home'), progressed through a rural setting, joined a UK two-lane dual-carriageway (comprising the majority of their journey experience) and finally arrived in an urban/city environment (described as their 'place of work'). SAE level-3 automation was only available on the dual-carriageway (and participants were made aware of this prior to taking part). All roads were populated with moderate to high levels of traffic (befitting a commute-style journey), authentic road signage and geo-typical roadside artefacts and terrain.

The vehicle's capabilities were described in detail to participants at the start of the study based on the SAE definition of level 3 automated control [4]. Specifically, participants were told within an information sheet:

> *"The vehicle is capable of controlling all aspects of the driving task. However, you may be required to resume manual control given appropriate notice."*

Procedure

Participants began by driving manually (i.e. they were responsible for all primary control actions). Automated control was only made available when the vehicle joined the dual carriageway, which occurred approximately 5 minutes into the journey: this was indicated to drivers via an internal HMI (see below for further details).

Participants requested the hand-over (and take-over) of control using a spoken command. Spoken language interfaces are increasingly commonplace in technological applications and can offer a quick, intuitive and increasingly reliable means of interaction. In addition, they do not require users to learn a new HMI or interaction method, but instead typically rely on the use of familiar, 'natural' language. In the current study, participants were required to use an appropriate command, preceded by the keyword 'AutoCar' (akin to current commercial products, such as Amazon Alexa). For example, "*AutoCar: start automated driving*" requested the hand-over of control from participant to car. This initiated the transfer of control, although in practice, it was manually triggered by the researcher.

During scheduled hand-overs (i.e. at the end of the dual carriageway, approximately 20 minutes later), drivers were presented with a 'prepare to drive' multi-modal warning (auditory and visual), delivered 60 seconds prior to takeover. This was followed by a takeover request ('resume control') delivered 10 seconds prior to the provision of control. Ten seconds was chosen for the duration of transfer of control based on our understanding of current technological capabilities, and corresponds with existing literature and recommendations [9].

Thereafter, participants completed the journey driving manually (a further 5 minutes, approximately). Completing the journey manually was intended to create an ecologically-valid full 'journey' experience, but also to ensure that participants' driving skills and exposure were 're-calibrated' before leaving the testing environment.

On day four (of five), drivers were presented with an emergency handover (due to inclement weather). In this situation, they only received the 10 seconds 'resume control' notification accompanied by an urgent alarm. For those participants who were provided with the feedback-HMI, the sensors 'failed' and turned red.

Measures

Following each drive, participants completed the trust in a specific automation questionnaire [10], the technology acceptance questionnaire [11], and the situational awareness rating scale (SART) [12]. In addition, manual driving performance data were collected using the simulation software. This was interrogated for the first ten seconds of manual driving, immediately after participants took over control, to explore their ability to safely resume control after the extended period of automation. Finally, four video cameras were strategically located within the vehicle to record participants' behaviour. The videos were also used to extract salient visual indicators (gaze directed towards road situation, mirror checks etc.) during the handover period and immediately thereafter (i.e. during the initial period of manual driving). Videos were coded and analysed on a frame-by-frame basis for the period(s) of interest using the BORIS event-logging software (http://www.boris.unito.it). Participants were thus in attendance for approximately 45-minutes per day (30-minutes driving experience and 10-15 minutes to complete questionnaires).

Human-Machine Interfaces

An in-vehicle HMI was installed in the centre console of the vehicle, providing both feedback and take-over advice.

The *Feedback HMI* provided details of system 'health' by displaying the car's sensors as green, amber or red (indicating increasing levels of severity) (Table 1). Drivers were notified of changes to sensor status (e.g. green to amber) with a non-intrusive tone, and the associated change of colour. This occurred seldom during the week and only for short periods of time (circa 30-seconds), without any accompanying external stimuli in the driving environment. The feedback HMI was intended to provide an 'intuitive' overview of the current state of the vehicle's sensors and control system, and therefore an indication of the vehicle's ability to safely provide control. In the real world, factors that could theoretically influence SAE level 3 automation might include the presence of an external hazard or a problem with the operational integrity of the sensors or control system, although the feedback HMI was not intended to replicate a specific real-world system.

In addition, a *Takeover HMI* provided guidance during hand-over requests, displaying either 'top-down' or 'bottom-up' hand-over advice based on driving skills hierarchy [7] (Table 2). 'Bottom-up' advice instructed participants to 'resume control' (i.e. at the operational/control level) – the assumption is that these drivers subsequently acquire 'tactical' knowledge after they had begun driving. In contrast, 'top-down' advice initially advised participants to 'check for hazards' ('tactical') – by guiding their visual attention towards the road-scene, mirrors etc. – before instructing them to resume control; in addition, participants receiving 'top-down' advice were provided with a count-down timer (displayed on the HMI) for each stage of the handover.

Design

The specific nature of the information displayed on the internal HMI varied between four groups of participants. Participants were either provided with feedback during

Representation	Status
	**Green**: The sensors are working fine. No action required
	**Amber**: Warning. The sensors may be faulty or dirty. No immediate action required
	**Red**: Sensors failure. The driver must immediately take control of the vehicle

Table 1. Automated system 'Feedback-HMI'.

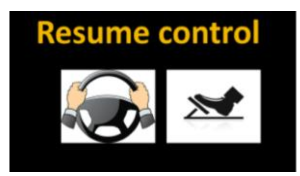
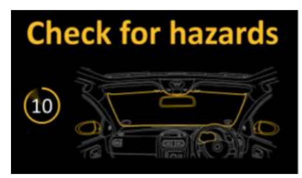
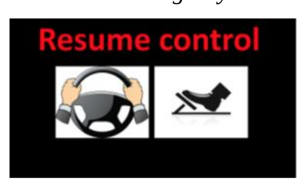
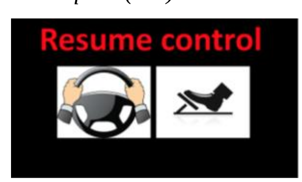
'Top-Down'	'Bottom-Up'
Routine Take-Over Request (60s)	
	
Routine Take-Over Request (10s)	
	 
Emergency Take-Over Request (10s)	
	

Table 2. Text-based and pictorial messages for 'top-down' (left) and 'bottom-up' (right) 'Takeover-HMI'. Changes were accompanied by an appropriate alerting tone and/or voice message.

periods of automation, or had no feedback (Table 1). In addition hand-over advice was presented either as 'top-down' or 'bottom-up' (Table 2). This resulted in four conditions that were delivered in a 2×2 between-subjects design. Each participant only experienced one of the four conditions, and this remained consistent throughout the week. For participants who did not receive the feedback shown in Table 1, the HMI simply indicated 'autonomous mode' or 'manual driving', throughout, as appropriate. System feedback and take-over advice was delivered via an interactive PowerPoint presentation controlled remotely by the researcher.

RESULTS AND ANALYSIS

The study aimed to explore how drivers' behaviour and attitudes developed over the week, and the impact of different HMI interventions. This was achieved by analysing the activities undertaken by participants, and their associated ratings of trust, technology acceptance and situational awareness, captured using recognised rating scales [10, 11, 12]. In addition, driving performance measures were recorded directly from the STISIM software and analysed for the first ten seconds of manual driving, immediately after participants took over control. Figures show within-subjects error bars, unless otherwise indicated.

Secondary Activities

Frame-by-frame video coding was undertaken to identify the types of devices and activities undertaken by participants during periods of automation. Behaviours were coded at a 'device' level – specifying the primary device used, rather than the specific activity undertaken using the device, for example, 'phone' is reported rather than 'sending text message'. This was to ensure the privacy of participants. To explore any behavioural adaptations across the week's experience, results from day one (Monday), day 3 (Wednesday) and day five (Friday) are presented here (Figure 2).

At an aggregate level, results are perhaps unsurprising. The most common item used by participants, for example, was their mobile phone. This was used by over eighty percent of participants, and remained consistent throughout the week. The next most popular activity was reading a book/magazine or printed papers (identified as 'Book'), with up to twenty-five percent of participants engaging in some form of physical reading activity during the week. Again, this remained largely unchanged as the week progressed. Other recorded activities were using a laptop ('PC') or tablet computer, and there was also some evidence of people sleeping, although this was rare. Many of these activities had a marked effect on drivers' state and physical surroundings, with participants repurposing their cabin or relaxing their seating position or posture to accommodate secondary task execution. On occasion, drivers had completely repurposed their cabin and were engaged in using multiple items (Figure 3), here shown at the point of unexpected, emergency handover.

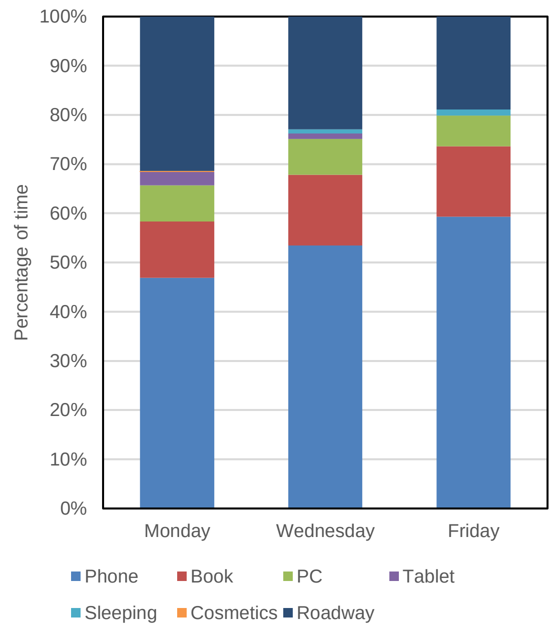


Figure 2. Mean percentage of time spent engaged in secondary activities and looking at roadway during automation

Figure 3. Participant engaged in multiple activities at point of unexpected, emergency handover request

Of note, is that most of these activities were highly engaging, generally with strong visual, cognitive and manual elements, and therefore had the potential to distract drivers. Moreover, participants appeared quite comfortable, from day one, to engage with these tasks – soon after the opportunity presented itself – despite their ongoing responsibilities at SAE level-3. The concern is that this potentially impacts on drivers' readiness to resume control, with activities contributing to visual distraction (e.g. reading a book), cognitive distraction (e.g. participant engrossed in a mobile phone exchange) and manual distraction, e.g. participant using a laptop. As a consequence, these drivers were not only required to discharge their secondary activity prior to resuming control, but also to re-establish appropriate driving posture (reposition seat, sit upright etc.). For some drivers, secondary activities also necessitated the wearing of reading

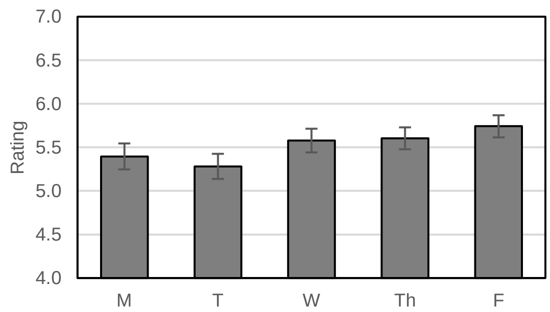


Figure 4. Trust ratings with standard error bars (1 = Strongly disagree, 7 = Strongly agree)

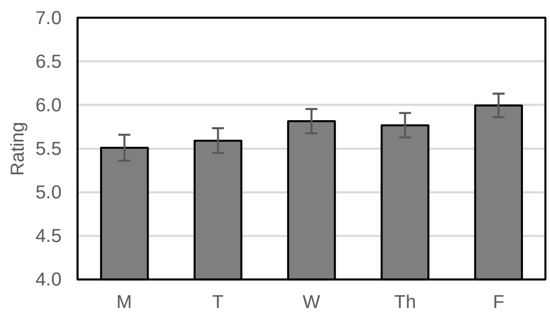


Figure 5. Technology acceptance ratings with standard error bars (1 = Strongly disagree, 7 = Strongly agree)

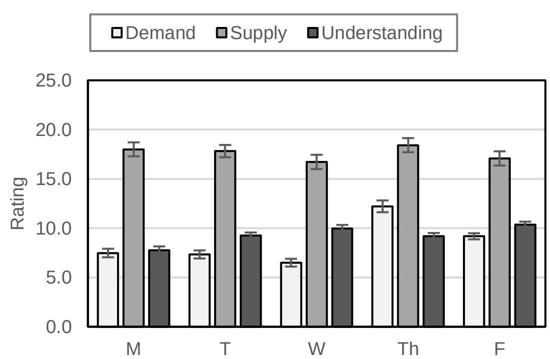


Figure 6. Situational Awareness (SART) ratings, with standard error bars

glasses, and again these needed to be removed and stored prior to taking control.

There was some evidence that even while participants were engaged in their chosen activities, they still returned their visual attention to the road-scene, through occasional glances. Figure 2 shows the mean percentage of time that participants spent engaged in (i.e. directing visual attention towards) their secondary activities and that directed to the roadway. Of particular note, is that on day one (Monday), almost 70% of participants' visual attention was directed at the secondary activities, with the majority of this directed towards smartphones. Even so, approximately a third of the time was spent looking at the roadway. However, there is a clear pattern as the week progresses, with participants directing less and less visual attention to the road-scene, and more at their secondary activity. By day 5 (Friday), over 80% of drivers' visual behaviour is directed at the secondary device, and less than 20% at the road-scene. Statistically, the amount of time spent looking out of the vehicle reduced significantly as the week progressed ($F(2,100) = 8.1$, $p = .001$, $\eta p^2 = .139$).

Trust

Responses to the trust in a specific technology scale [10] show that initial trust (on day one) was above the scale median of four (rating: 5.4), and thereafter, trust generally increased throughout the week (Figure 4). Interestingly, there was no apparent detriment to trust following the emergency handover, either on day four, either immediately following the experience or indeed, on the next day. A repeated-measures ANOVA revealed that differences between daily ratings were significant ($F(4,196) = 6.54$, $p < .001$, $\eta p^2 = .118$). Pairwise comparisons with Bonferroni corrections indicated that ratings on days three, four and five (Wed, Thurs and Fri) were significantly higher than those made on day two (Tues) ($p < .001$, $= .047$ and $.001$, respectively).

Technology Acceptance

Ratings for technology acceptance [11] indicate that initial acceptance (on day one) was again above the scale median of four (rating: 5.5), and generally increased throughout the week (Figure 5), although there was an apparent fall in acceptance, following the emergency handover experienced on day four.

A repeated-measures ANOVA revealed that differences between daily ratings of acceptance were significant ($F(4,196) = 11.02$, $p < .001$, $\eta p^2 = .184$). Pairwise comparisons with Bonferroni corrections showed that ratings of acceptance on day five were higher than those made throughout the week. In addition, ratings on day three were higher than those made on days one and two. This indicates that there was in fact no detriment to acceptance following the emergency handover on day four. Moreover, acceptance actually increased the following day.

Situational Awareness

Situational awareness (SA) was determined using the situational awareness rating technique (SART) [12]. Ratings were combined to elucidate participants' evaluation of so-called 3D-SART (Figure 6), incorporating:

1. The demands on their attentional resources (complexity, variability, and instability of the situation) – *Demand*
2. The supply of attentional resources (division of attention, arousal, concentration, and spare mental capacity) – *Supply*

3. Their understanding of the situation (information quantity, and information quality) – *Understanding*

The demand on attentional resources (demand) initially dropped day-on-day as the week progresses (from day one to day three). However, there was a sharp increase on day four – understandably, given the emergency handover. Although there was a reduction on the following day (day five), ratings remained higher than they were earlier in the week, suggesting a residual effect of the day four emergency. A repeated-measures ANOVA revealed that differences between daily ratings of demands on attentional resources were significant ($F(4,196) = 33.5$, $p < .001$, $\eta p^2 = .406$). Pairwise comparisons, with Bonferroni corrections for multiple comparisons, showed that ratings on day four were significantly higher than all other days ($p_{max} = .001$), and ratings on day five were significantly lower than day four ($p = .001$), but remained significantly higher than days one, two and three ($p = .020, .003, < .001$, respectively).

A similar pattern emerges for the supply of attentional resources (supply), with a decline from day one to day three, a peak on day four and a drop on day five. However, although a significant effect was revealed ($F(4,196) = 3.40$, $p = .010$, $\eta p^2 = .065$), differences were only significant between day three and day four ($p = .020$), indicating a marked effect of the emergency handover on day four.

In contrast to other situational awareness measures, participants' understanding of the situation (understanding) increased as the week progressed. There was evidently some 'confusion' following the unexpected, emergency handover on day four, indicated by a drop in ratings on this day, although participants' understanding appeared to increase again the following day. Differences were indeed significant ($F(4,188) = 13.89$, $p < .001$, $\eta p^2 = .228$), with pairwise comparisons (with Bonferroni corrections) indicating that day one ratings were significantly lower than the rest of the week ($p_{max} = .018$). In addition, ratings on day five were significantly higher than those made on day four ($p = .002$).

Driving Performance

In line with previous, similar investigations (e.g. [13]), driving performance during the first ten seconds immediately after resuming manual control was analysed by dividing this into one-second time intervals. Data were captured directly by the simulation software and interrogated to explore variability in lateral behaviour (lane position) (absolute and standard deviation), as a key indicator of vehicle control.

The mean lateral lane position from the lane centre, immediately following the take-over of control (represented as the position of the centre of vehicle) is shown in Figure 7. There was a clear tendency for drivers to move to the left (away from oncoming traffic – the study was conducted in the UK) following the resumption of control. Driving performance data show that the lateral movement immediately after resuming control was highly variable. The highest deviation of absolute lane position (from lane centre)

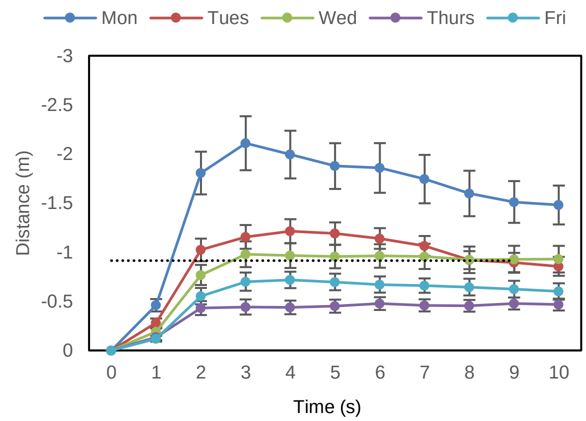


Figure 7. Mean position of centre of vehicle in 1s intervals after resuming manual control, with standard error bars. Note: '0' distance represents lane centre, and negative distances are to left.

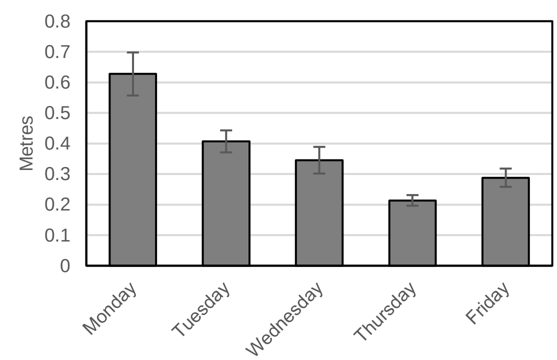


Figure 8. Mean standard deviation of lane position (SDLP) during the 10s after resuming manual control, with standard error bars

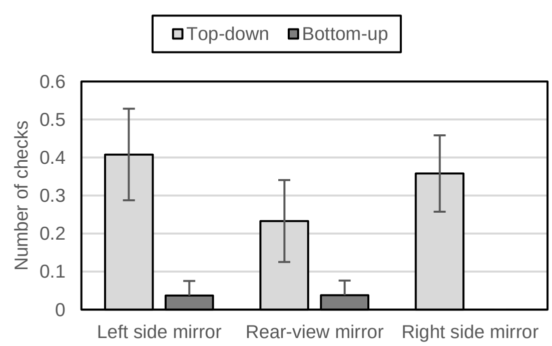


Figure 9. Mean number of mirror checks during top-down and bottom-up advice (with standard error bars)

was evident on day one, with drivers, on average, moving up to over 2 metres to the left from the lane centre (peaking at 3-seconds after resuming control). On days one, two and three the vehicle actually crossed the lane demarcation and moved into the adjoining ('inside') lane (the dotted line in

Figure 7 shows the location of the centre of the vehicle at the point at which the edge of the vehicle would cross the lane boundary). It is also notable that the level of lateral control demonstrated by drivers tended to improve as the week progressed. On day one, drivers typically did not manage to regain their lane position even after 10-seconds of manual driving, remaining notably approximately 1.5-metres outside of their lane. On days two and three, it took drivers 8-seconds on average to recover their vehicle to within the lane limits. On days four and five, drivers managed to keep their vehicle within the lane boundaries, though still with notable (approximately 0.5 to 0.8m) lateral deviation. Interestingly, the smallest deviation form lane centre was evident on day four, after the emergency handover.

The standard deviation of lane position (SDLP) is indicative of unstable 'wavering' behaviour. Considered over the entire 10-second period, it appears that SDLP was highest on day one, and thereafter apparently improves as the week progresses (Figure 8). Moreover, the lowest SDLP appears to occur on day four. A Shapiro-Wilk test revealed that the SDLP data was not normally distributed ($p < .001$). Consequently, aligned rank transform was used to analyse these data. This revealed a significant effect ($F(4,235) = 13.69$, $p < .001$, $\eta p^2 = .19$). Pairwise comparisons with Bonferroni corrections, showed that SDLP on day one was higher than days three, four and five ($p_{max} = .002$). Day two was also significantly higher than days four and five ($p_{max} = .028$), and day three was higher than day four ($p = .048$). Whilst confirming that SDLP on day four was indeed significantly lower than days one, two and three, results show no significant difference for SDLP between days four and five ($p = .374$).

HMI Interventions

A further aim of the study was to explore the effect of novel HMIs to help keep drivers 'in-the-loop' during periods of automation using a *Feedback-HMI* (see: Table 1), and to improve their performance during take-overs with a *Takeover-HMI* (see: Table 2). To explore these effects, the emergency handover on day four was considered in isolation. Selecting day four for analysis ensured that participants had had sufficient time to become accustomed to the experience. Moreover, the unexpected, emergency situation (inclement weather necessitating the hand-over of control) provided the perfect opportunity to 'stress test' the HMI interventions. Further video coding was therefore conducted to classify participants' mirror-checking behaviour during the take-over request (i.e. following the delivery of the 'take control' instruction but before drivers actually resumed manual control). This provides an indication of participants' attempts to re-engage themselves with the driving scene (i.e. rebuild their situational awareness). In addition, 'driver readiness' is defined as the time at which participants made their first glance to the road scene *and* had at least one hand on the steering wheel, in line with similar research [8].

Mirror Checks

A two-way Multivariate Analysis of Variance (MANOVA) showed a significant main effect of *Takeover-HMI* ('top-down' versus 'bottom-up') on checks to right side, left side and rear-view mirrors ($F(3, 43) = 4.82$, $p = .006$, Wilks' $\Lambda = .748$, $\eta p^2 = .252$) (Figure 9). There was no main effect of *Feedback-HMI* ($F(3, 43) = .278$, $p = .841$, Wilks' $\Lambda = .949$, $\eta p^2 = .019$), and no significant interaction effect of *Takeover-HMI* and *Feedback-HMI* on the frequency of mirror checks, $F(3, 43) = .852$, $p = .473$, Wilks' $\Lambda = .94$ $\eta p^2 = .056$. A subsequent one-way ANOVA revealed that significantly more checks were made to the right and left side mirrors when drivers were provided with 'top-down' (*Takeover-HMI*) guidance. No differences in checks to the rear-view mirror were found for *Takeover-HMI*.

In contrast, no checks were made to the right-side mirror by drivers who received 'bottom-up' (*Takeover-HMI*) advice, and only 3.7% of these drivers checked their left-side mirror (compared to 36.4% of 'top down' drivers checking both left and right) (although no differences were evident between conditions in the number of checks to the rear-view mirror).

Time to 'Driver Readiness'

A two-way ANOVA revealed that there was a significant main effect of *Feedback-HMI* during automation on the time to 'driver readiness' ($F(1, 45) = 12.71$, $p = .001$, $\eta p^2 = .221$), reducing this by 2.1-seconds compared to situations in which no feedback was provided.

DISCUSSION

The experience provided by future, automated vehicles is expected to be fundamentally different to manual driving [6]. As such, the role and responsibilities of the driver will change. Understanding how drivers will behave – particularly at intermediate levels of automation, where vehicle control responsibilities may be shared – and the influence that this has on their ability to resume manual control (if required) is therefore important. The study therefore employed a novel, longitudinal driving simulator study – providing multiple exposures to SAE level 3 automation – to capture behavioural adaptations and explore the effect of different HMI interventions to support drivers.

Whereas longitudinal driving studies have been employed previously (e.g. [14, 15, 16]), these have tended to consider automated control technologies, such as adaptive cruise control (ACC), in isolation and under strict experimental protocols. In contrast, this is the first longitudinal simulator study to expose drivers to a full journey experience at SAE level-3 automation with multiple, repeated visits, and provide them with the freedom to exhibit behaviours that were most natural to them.

Behavioural Adaptations

Results show that when provided with the opportunity to delegate control to the vehicle, drivers quickly availed themselves and undertook a range of self-selected activities – typically with high visual, manual and cognitive demands. While the types of devices and activities did not change significantly during the week, the proportion of time (determined by the visual attention directed towards them) increased from approximately 70% on day one to over 80% on day five (i.e. less than 20% was directed at the road-scene).

Ratings of trust and technology acceptance were high from day one. Moreover, trust and technology acceptance ratings generally increased as the week progressed (similar trends were seen associated with ACC by Beggiato and Krems [15] and Kazi et al. [14]). Of particular note here, was that there was no apparent detriment to trust or technology acceptance after the emergency handover event on day four – either immediately thereafter, or on the next day. In fact, ratings of trust were significantly higher at the end of the week compared to the start, and technology acceptance was highest on day five. This is important as the level of trust and acceptance that users place in technology is expected to influence their uptake of the technology [17] – if it is not seen as 'acceptable' by users, they will not buy it and even if they do, they may disable it out of frustration (disuse) (if/where possible), or use it in a manner unintended by designers (abuse) [18]. Moreover, from a methodological point of view, differences in ratings between day one and day five (as seen here) could lead to different conclusions if considered in isolation.

Ratings of situational awareness revealed that for the first part of the week (days one to three), participants tended to report an ongoing reduction day-on-day in the demands-on and supply-of their attentional resources (division of attention, arousal, concentration, and spare mental capacity) and an increase in their understanding of the situation. This is likely to be because they became more familiar with their experience of the automated vehicle – rather than their awareness of the driving situation *per se* – although it is worth considering that the effect may also be interpreted from the perspective of increased familiarity with the simulated driving experience and experimental conditions. However, the unexpected, emergency handover on day four had a marked effect – significantly increasing the SART-demand placed on drivers' attentional resources, and reducing their understanding of the situation. While this is to be expected, given the nature of the perturbation, it is particularly interesting to observe that the ratings of situational awareness returned to their previous magnitude the day after, indicating no residual effects from the emergency handover situation.

Unsurprisingly, drivers' secondary behaviours had a significant impact on driving performance and behaviour, with data showing that lateral movement (both absolute and standard deviation of lane position) was highly variable immediately after resuming control, and remained so for up to 10 seconds thereafter. Perhaps more interesting is that lateral displacement was always to the near-side (left), suggesting that drivers inherently determined this to be a safer course of action (rather than steering towards on-coming traffic). It is worth highlighting that during the study, the vehicle operated in automated mode in lane two (the 'outside' lane) of the dual carriageway (presented to participants as a dedicated 'autonomous' driving lane). Thus, when automated control was relinquished, drivers could choose whether to remain in lane two, or move to the left into lane one. While the general tendency to 'drift left' might suggest that drivers were actually choosing to move into lane one, the fact that they ultimately attempted to regain their central position in lane two suggests that this was not in fact their intended strategy, but rather that they required time to readapt or re-calibrate primary control mechanisms.

Even so, lateral control improved as the week progressed, suggesting that drivers were able to develop their own strategies to improve immediate take-over performance (such strategies could potentially be incorporated within future training and best practices). Interestingly, the lowest levels of lateral displacement were evident immediately after the emergency handover on day four. Initially counter-intuitive, we suspect that this is an indicative of heightened arousal – and an associated increase in focus of attention (mental effort and mobilisation of resources) – due to the nature of the emergency handover (in particular, the urgent alarm). This is also supported by the SART ratings, which showed an increase in the supply of attentional resources – in particular – on this day.

HMI Interventions

A further aim of the study was to explore possible HMI interventions to help keep drivers 'in-the-loop' during periods of automation and to rebuild their situational awareness prior to the hand-over of control. There were some notable differences in mirror-checking behaviour associated with the different 'top-down'/'bottom-up' Takeover-HMIs – something that is considered essential for safe driving and helps drivers to maintain situational awareness [19]. Indeed, providing drivers with 'top down' guidance (encouraging them to check for hazards) during the 10-second transition led to significantly more checks being made to the right and left-side mirrors, even during the unexpected, emergency hand-over. This suggests that such strategies could be employed to improve drivers' situational awareness prior to providing control even during a relatively short time-frame – although it did not necessarily translate to improved driving performance immediately after the hand-over. However, not all drivers who received 'top-down' advice during the study actually checked their mirrors, suggesting that prompting participants to check for hazards may not be a fully effective method to build situational awareness during handover. Even so, no lane change was required as part of the emergency handover, and surrounding traffic was sparse at the point of hand-over. It is therefore feasible that drivers' mirror-checking behaviour may have been influenced by these factors – their behaviour may have been different had they been required to undertake a lane-change manoeuvre immediately after resuming control, or traffic density was higher – factors that could be explored in future work.

It is also noteworthy that those participants who were provided with the Feedback-HMI demonstrated no difference in mirror-checking behaviour during the handover, suggesting that keeping drivers ‘in-the-loop’ during automation had no impact on their behaviour during the transfer of control, and further supporting the need for the provision of additional (‘top-down’) information or guidance during the handover. Nevertheless, it is possible that drivers expected the Feedback-HMI to highlight potential obstacles at the point of handover, suggesting overreliance and potential errors of omission, i.e. drivers failed to implement actions because they are not instructed to do so by the vehicle (as also observed by Eriksson et al [20]).

Limitations

It is important to recognise that the research was conducted at SAE ‘level 3’ automation, in so far participants were told that they may be required to resume manual control during periods of automation, given appropriate notice. Nevertheless, no restrictions were applied to the type of activities that drivers could undertake (and consequently, secondary devices and activities were not controlled across groups). In practice, this may seem somewhat inconsistent with our understanding of SAE level 3 automation, in that drivers may still be required to take-over control at some point. It would therefore seem prudent that a future SAE level 3 vehicle would not permit certain types of activities, and warn, penalise or attempt to re-engage the driver accordingly. However, we did not want to prejudice their choices.

In addition, the range of secondary activities undertaken may have influenced some of our measures. For example, participants undertaking activities that had high visual, manual and cognitive elements (e.g. working on a laptop placed in front of the steering wheel), may have taken considerably longer to detach from their non-driving related activities than participants selecting less demanding activities, for example, those casually glancing at their smartphone at the point of hand-over. Thus, it is feasible that reaction times may be influenced by the secondary device being used, although in practice, the range of devices used and activities observed during the study was fairly limited.

Overall, the intention in conducting the research in this manner was to explore the type of activities that drivers would expect to undertake in a vehicle offering SAE level 3 capabilities (defined to them at the outset) – during a daily commute journey. Consequently, we provided drivers with the freedom to exhibit behaviours that were most natural to them – which may therefore arise whether permitted or not within a vehicle. This also detracted from the experience being seen as a ‘controlled’ experiment (in which participants were only able to select from a limited range of activities), and avoided pre-empting any legislative control.

Finally, it is worth highlighting that the results as presented relate to the specific set of controlled conditions to which participants were exposed (i.e. in a medium-fidelity driving simulator, with specific experimental HMIs etc.). Whilst every effort has been made to ensure that these factors are as realistic as possible (based on our understanding and expectations of future vehicles), there are aspects of the experimental set-up that were chosen for experimental convenience (though notably informed by expert opinion) that may be more complex in the real world. For example, in practice, it is debatable whether an automated vehicle would have sufficient self-awareness to be able to show the current status of its sensors to its passengers with the degree of clarity that we did. Moreover, there are unavoidable limitations when conducting such research in a driving simulator, associated with the fidelity of experience, low risk perception etc. As such, caution should be applied when drawing conclusions regarding ‘absolute’ behaviour.

CONCLUSION

Results as a whole suggest clear benefits associated novel HMI designs to keep drivers in-the-loop and improve take-over performance. It is evident that such interventions can influence drivers’ behaviour during and immediately following transitions of control, although further work is required (for example, to explore different HMI designs and strategies). In addition, the study demonstrates a clear need to provide multiple, repeated exposures during the evaluation of future, immersive technologies (such as full-vehicle automation) to ensure that behavioural adaptations are exposed. Indeed, it is possible that different conclusions could be drawn if taking results from day one (or indeed, from day five) in isolation.

ACKNOWLEDGEMENTS

The research was funded by the RAC Foundation and the authors would like to thank Elizabeth Box and Anneka Lawson for their valuable guidance and support. We would also like to thank Hannah White and Emily Shaw for their assistance during analysis.

REFERENCES

1. N. Merat, A. H. Jamson, F. C. H. Lai and O. Carsten, “Highly Automated Driving, Secondary Task Performance, and Driver State,” *Human Factors*, vol. 54, no. 5, pp. 762-771, 2012.

2. D. R. Large, G. E. Burnett, A. Morris, A. Muthumani and R. Matthias, “A Longitudinal Simulator Study to Explore Drivers’ Behaviour during Highly- Automated Driving,” in *Proceedings of the Advances in Human Aspects of Transportation*, New York, NY, 2018.

3. Department for Transport (DfT), “The Pathway to Driverless Cars: Summary Report and Action Plan,” DfT Publications, London, 2015.

4. Society of Automotive Engineers (SAE), “Taxonomy and Definitions for Terms Related to On-Road Motor

Vehicle Automated Driving Systems (Standard No. J3016)," 2014.

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

6. M. Kyriakidis, J. de Winter, N. Stanton, T. Bellet, B. van Arem, K. Brookhuis, M. Martens, K. Bengler, J. Andersson, N. Merat and N. Reed, "A human factors perspective on automated driving.," *Theoretical Issues in Ergonomics Science,* pp. 1-27, 2017.

7. J. Michon, "A critical view of driver behavior models: what do we know, what should we do?," *Human behavior and traffic safety,* pp. 485-524, 1985.

8. K. Zeeb, A. Buchner and M. Schrauf, "What Determines the Take-Over Time? An Integrated Model Approach of Driver Take-Over After Automated Driving," *Accident Analysis and Prevention,* vol. 78, pp. 212-221, 2015.

9. V. Melcher, S. Ruah, F. Diederics, H. Widlroither and W. Bauer, "Take-Over Requests for Automated Driving," *Procedia Manufacturing,* vol. 3, pp. 2867-2873, 2015.

10. D. Mcknight, M. Carter, J. Thatcher and P. Clay, "Trust in a specific technology: An investigation of its components and measures.," *ACM Transactions on Management Information Systems (TMIS),* vol. 2, no. 2, p. 12, 2011.

11. V. Venkatesh and F. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies.," *Management science,* vol. 46, no. 2, pp. 186-204, 2000.

12. R. Taylor, "Situational Awareness Rating Technique (SART): The development of a tool for aircrew systems design. Situational Awareness in Aerospace Operations (AGARD-CP-478)," Neuilly Sur Seine, France: NATO-AGARD, 1990.

13. S. Pampel, D. Large, G. Burnett, R. Matthias, S. Thompson and L. Skrypchuk, "Getting the driver back into the loop: the quality of manual vehicle control following long and short non-critical transfer-of-control requests," *Theoretical Issues in Ergonomics Science,* pp. 1-19, 2019.

14. T. Kazi, N. Stanton, G. Walker and M. Young, "Designer driving: drivers' conceptual models and level of trust in adaptive cruise control," *International Journal of Vehicle Design,* vol. 35, no. 3, pp. 339-360, 2007.

15. M. Beggiato and J. Krems, "The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information," *Transportation research part F: traffic psychology and behaviour,* vol. 18, pp. 47-57, 2013.

16. M. Beggiato, M. Pereira, T. Petzoldt and J. Krems, "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, 2015.

17. M. Ghazizadeh, J. Lee and L. Boyle, " Extending the Technology Acceptance Model to assess automation.," *Cognition, Technology & Work,* vol. 14, no. 1, pp. 39-49, 2012.

18. R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human factors,* vol. 39, no. 2, pp. 230-253, 1997.

19. N. Li and C. Busso, "Driver Mirror-Checking Action Detection Using Multi-Modal Signals.," in *Proceedings of the 6th Biennial Workshop in Digital Signal Processing for In-Vehicle Systems and Safety,* Red Hook, NY, 2013.

20. A. Eriksson, S. M. Petermeijer, M. Zimmermann, J. C. F. De Winter, K. Bengler and N. A. Stanton, "Rolling Out the Red (And Green) Carpet: Supporting Driver Decision Making in Automation-to-Manual Transitions," *IEEE Transactions on Human Machine Systems,* 2018.