Article:

Louw, T orcid.org/0000-0001-6577-6369, Kuo, J, Romano, R et al. (3 more authors) (2019) Engaging in NDRTs affects drivers' responses and glance patterns after silent automation failures. Transportation Research Part F: Traffic Psychology and Behaviour, 62. pp. 870-882. ISSN 1369-8478

https://doi.org/10.1016/j.trf.2019.03.020

Engaging in NDRTs affects drivers' responses and glance patterns after silent automation failures

Tyron Louw[1], Jonny Kuo[2], Richard Romano[1], Vishnu Radhakrishnan[1], Michael G Lenné[2], Natasha Merat[1]

[1]Institute for Transport Studies, University of Leeds, UK

[2]Seeing Machines, Canberra, Australia

Abstract

The aim of this study was to understand driver responses to "silent" failures in automated driving, where automation failed during a simulator drive, without a take-over warning. The effect of a visual non-driving related task (NDRT) and a road-based vigilance task presented drivers' take-over response and visual attention was also investigated. Currently, automated driving systems face a number of limitations that require control to be handed back to the driver. Much of the research to date has focused on explicit take-over requests (ToRs) and shows that drivers struggle to resume control safely, exacerbated by disengagement from the driving task, for instance, due to the presence of NDRTs. However, little is known about whether, and how, drivers will respond to more subtle automation failures that come without a warning, and how this is affected by NDRT engagement. Thirty participants drove a simulated automated drive in two conditions, which had 6 silent automation failures each (3 on a Curve, 3 in a Straight), with no ToRs. In one condition, drivers were required to constantly monitor the road, which was enforced by a road-based vigilance task (VMS Only). In the other, drivers performed an additional visual NDRT, requiring them to divide their attention (VMS + Arrows). Results showed that, in both conditions, all drivers eventually detected and responded to all silent automation failures. However, engaging in an additional NDRT during automation resulted in significantly more lane excursions and longer take-over times. Adding a visual NDRT not only changed the distribution of drivers' visual attention before and after the failure but also how they divided their attention between information on the road environment and the human-machine interface, which provided information on automation status. These results provide support for how driver monitoring systems may be used to detect drivers' visual attention to the driving task and surroundings, and used as a tool for encouraging driver intervention, when required.

Keywords: automated driving; eye-tracking; take-over; distraction; driver behaviour; automation failure

1. Introduction

Semi-autonomous driving systems have entered the commercial market in recent years. At a basic level, these systems combine Adaptive Cruise Control (ACC) with a Lane Keeping System (LKS), which manage the vehicle's longitudinal acceleration to maintain a desired distance from a lead vehicle, and control the vehicles steering to maintain its position in a lane, respectively. For the most part, the benefit of these types of systems is comfort-related, though they do offer some other benefits. For example, ACC could be used to avoid violating speed limits and improving fuel efficiency (Benmimoun, Pütz, Zlocki, & Eckstein, 2012), while LKS could be used to avoid unintended lane departures.

Despite the apparent benefits of these types of automated driving systems, there remains a major drawback: they still require human involvement in the driving task, both in terms of supervision, and also intervention (Merat et al., 2018). Such systems are widely referred to as Level 2 (L2) or Level 3 (L3; SAE, 2016) systems, and both require drivers to always be ready to take control of the dynamic driving task (DDT), following a "take-over request" (ToR). However, whereas L2 systems require drivers to constantly monitor the environment, L3 systems allow drivers to look away from the road environment, only requiring a response to a "request to intervene".

Despite these requirements, recent driving simulator studies have shown that drivers' visual attention to the road centre is much reduced during SAE L2 and L3 driving (SAE, 2016), compared to when they are in manual control of the vehicle (Louw, Merat, & Madigan, 2015; Zeeb, Buchner, & Schrauf, 2016), mostly because drivers tend to look around more during such automation engagement. This reduced visual attention to the road centre is further diminished during driver engagement in other (visual) non-driving-related tasks (NDRTs), which may be voluntary (e.g. Carsten et al., 2012; Llaneras, Salinger, & Green, 2013) or enforced by the experimental conditions (e.g. Louw, Madigan, Carsten, & Merat, 2017). Recent real-world observations of drivers in conventional vehicles also suggest higher incidents of driver distraction during automation, with engagement with mobile telephones and satellite navigation systems being particularly prevalent (Huisingh et al., 2015). As the degree of automation in vehicles increases, drivers' engagement with such distracting tasks is likely to increase (Naujoks, Purucker, & Neukum, 2016), perhaps to relieve boredom, or due to driver complacency, and a high trust in the automated system's capabilities (Banks et al., 2018), giving drivers the impression that engaging in NDRTs is safe. However, for safety reasons, drivers may occasionally glance towards the forward roadway, or familiarise themselves with the in-vehicle HMI.

Results from both driving simulator and real-world studies of Level 2 driving illustrate that reduced visual attention to the road centre, exacerbated by engagement in visual NDRTs, can be catastrophic, with poorer responses to critical incidents if drivers are required to take-over from the automated system, for instance, to avoid colliding with a lead vehicle (Louw, Madigan, Carsten, & Merat, 2017; Louw, Markkula, Boer, Madigan, Carsten, & Merat, 2017; Endsley, 2017; Banks et al., 2018). For example, in March 2018, an Uber test-vehicle collided with, and killed, a pedestrian that was crossing the road in its path (Marshall & Davies, 2018). The software used to interpret data from the vehicle's sensors misclassified the pedestrian as a "false positive", and, thus, determined there was no need to take evasive action, until the second before impact. The safety driver was responsible for monitoring the environment and initiating an emergency braking manoeuvre, required to mitigate a collision. However, because in the moments before impact the safety driver was engaged in an NDRT (watching a show on a mobile phone) and not monitoring the road, they did not notice the pedestrian and, therefore, were not able to detect and respond appropriately to avoid the collision.

To ensure humans remain sufficiently engaged with the driving task, perhaps also discouraging engagement in NDRTs, L2 vehicles currently available on the market are equipped with features that encourage regular contact with the steering wheel, for example by activating pressure sensors on the steering wheel, or providing sustained steering input (Pilot Assist II; Volvo, 2018). Rather than relying on steering inputs alone, some manufacturers have also incorporated camera-based sensing to assess and ensure driver attention to the driving environment (Cadillac, 2018). If drivers do not adhere to these requirements, the automated system disengages. However, some other implementations do not disengage the system when drivers are required to put their hands on the steering wheel, despite repeated visual and auditory alerts, for example, the previous version of Tesla's Autopilot.

As vehicles move from offering SAE L2 to L3 automation, drivers will no longer be obliged to monitor the driving environment, and will only be asked to respond to a ToR. As highlighted above, a growing number of studies are demonstrating that drivers face difficulties responding safely to ToRs. There are a number of possible contributing factors here, including drivers' over-reliance on, or errant mental model of, the system (See Victor et al., 2018), but also their inadequate perceptual-motor priming (Russell et al., 2017; Louw, Markkula, Boer, Madigan, Carsten, & Merat, 2017) or degraded situation awareness (Louw, Madigan, Carsten, & Merat, 2016). In simple terms, some drivers may not know exactly who is responsible, what they need to do, or how to do it.

Questions also remain about the extent to which increasing levels of automation (and different engagement strategies encouraged by in-vehicle HMI) will affect driver engagement during automation, and, therefore, their ability to detect and respond to more subtle automation failures. These are typified by insufficient, silent, or no warnings from the system, because it might not detect an issue or its limitations. For example, and to use another Tesla crash case (Lambert, 2018), a lane keeping system may track poorly marked or older lane boundaries, and take the vehicle off the intended lane path. In this case, the system may not recognise its error and, therefore, would not alert the driver. In such situations, the driver is ultimately responsible for detecting and responding to such errors and limitations. Flemisch et al. (2017) effectively use the concept of the Uncanny Valley to describe how, even in cases where drivers recognise such risks, the likely irregularity of automation failures will induce in drivers overtrust/overreliance or automation bias/complacency, resulting in drivers whose capability, whether cognitive or behavioural, cannot meet the requirements to rebalance and/or reduce the risk.

The majority of previous research has focused on how drivers engaged in different NDRTs during automation (e.g. Naujoks, Purucker, & Neukum, 2016) for different durations (e.g. Feldhütter, Gold, Schneider, & Bengler, 2017) respond to explicit take-over requests (ToRs; e.g. Louw, Merat, & Jamson, 2015), in different traffic situations (Madigan, Louw & Merat; 2018). Studies have also considered driver behaviour when there is a request to monitor the system and respond in the face of critical (e.g. Louw, Markkula, Boer, Madigan, Carsten, & Merat, 2017) and non-critical traffic scenarios (e.g. Eriksson & Stanton, 2017). The critical scenarios used often attempt to reflect actual, or anticipated, system limitations or failures, that would give rise to ToRs, for example, missing lane-markings (Melcher, Rauh Diederichs, Widlroither, & Bauer, 2015), inclement weather (Li, Blythe, Guo, & Namedo, 2018), approaching a construction zone (Naujoks, Mai, & Nuekum, 2014), or an obstacle in the forward roadway (Louw, Merat, & Jamson, 2015).

The impact on safety has primarily been measured as take-over performance, which has relied heavily on 1) reaction time measures, such as time to hands on the wheel, or time to first steer or brake, and 2) the subsequent lateral and longitudinal vehicle controllability measures, such as maximum and minimum accelerations, lateral deviations, lane departures, steering wheel reversals, and minimum time to collision (minTTC). However, the ability of each measure to reflect take-over *quality* is context-dependent. For example, lane departures would be more useful than minTTC in situations where there are no obstacles to collide with. Such approaches allow one to draw broad conclusions about the drivers' cognitive and physical abilities in a range of take-over situations. However, to ensure that drivers have at least relatively safe interactions with automated driving systems, we need to be able to predict behaviour on an individual level, to know whether and how they should be supported in the take-over. Few studies have investigated whether drivers can detect and respond to subtle automated driving failures, and how this is affected by drivers' attention to an NDRT.

1.1. Study Objectives

The aim of the present study was to observe driver behaviour during an SAE Level 2 automated drive, and investigate the effects of engagement in a visual NDRT and a road-based vigilance task, using glance behaviour to establish drivers' detection of silent failures of the automated system. The study reported here required drivers to read on-road variable message signs, to create an L2 driving environment where the driver was required to monitor the forward roadway. In this paper, we investigate the following research questions:

i. Can drivers detect and respond to a subtle failure of the automated system, cued by subtle changes in the proprioceptive feedback from the vehicle, and in the absence of any discernible vehicle-based warnings and associated HMI?
ii. How is this detection affected by drivers' engagement in a visual NDRT during automation?
iii. How does the road environment (Straight vs Curved road sections) affect failure detection?
iv. What is the pattern of drivers' glances before, and after, automation failure, and how is this affected by engagement in an NDRT?

2. Methodology

2.1. Participants

Following approval by the University of Leeds Ethics committee (Ref LTSEE-074), thirty regular drivers (19 male, mean age = 42 years ± 17; Mean driving experience: 22 years ± 16) were recruited for the study, using the University of Leeds Driving Simulator (UoLDS) database.

2.2. Materials

The experiment was conducted in the fully motion-based UoLDS, which consists of a Jaguar S-type cab housed in a 4m spherical projection dome with a 300° field-of-view projection system. The simulator also incorporates an 8 degree-of-freedom electrical motion system. This consists of a

500mm stroke-length hexapod motion platform, carrying the 2.5t payload of the dome and vehicle cab combination, and allowing movement in all six orthogonal degrees-of-freedom of the Cartesian inertial frame. Additionally, the platform is mounted on a railed gantry that allows a further 5m of effective travel in surge and sway.

To understand how drivers divided their attention between the road and NDRT, and establish when they noticed the silent failures, their eye-movements were recorded using a Seeing Machines Driver Monitoring System (DMS), which was placed on the dashboard above the HMI, as shown in Figure 1.

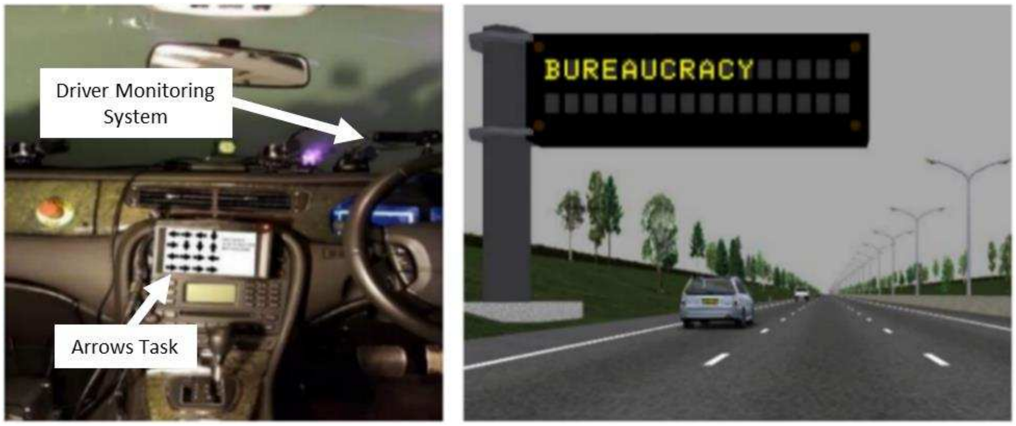

Figure 1 - Example of the Arrows NDRT (left), the position of the task in the vehicle (middle) and the VMS (right).

2.3. Design and Procedure

A within-participant design was used in the study with the factors being Non-Driving Related Task (NDRT: Variable Message Signs [VMS] Only, VMS + Arrows), Road Curvature (Curve, Straight), and Failure Order (1st, 2nd, 3rd). Participants were provided with written and verbal information about the study, and those not familiar with the simulator controls were shown how to use the simulator and NDRT task.

Following a 15-minute familiarisation drive, where the experimenter accompanied participants in the simulator, drivers completed two experimental drives (VMS Only, VMS + Arrows, described below, and presented in counterbalanced order). Both experimental drives were on a 3-lane UK motorway and began with a short manual drive (Approximately 14 km long, lasting ~8 minutes), which included a period of engagement in the Arrows NDRT. Following this manual drive, automation was engaged by drivers, which was available when the vehicle was placed in the centre of the middle lane, and if participants were driving within the designated speed limit (70 mph). A short break was available between the two drives, which required participants to leave the simulator dome and stretch their legs, before returning for the next drive.

During the VMS Only drive, drivers were asked to monitor the road environment and read aloud a series of low-frequency words presented on the VMS that was located on the side of the road (see Figure 1, for an example). In order to prevent a link between VMS presence and automation failure (see below), words were not displayed on all VMS, but this was not obvious to drivers, and so they continued to check the VMS until they passed it. Where the VMS did come on, the words appeared when the vehicle was 250 m away, and the message was clearly legible at this point.

During the VMS + Arrows drive, as well as reading the VMS words, drivers were required to complete the Arrows task (Jamson & Merat, 2004; Lee et al., 2013; see Figure 1), which required detecting and touching an upward facing Arrow, present in a 4x4 grid of Arrows on a touch-sensitive screen. This was a driver-paced task, which required identification of as many upward-facing Arrows as possible, with each detection prompting the presentation of the next 4x4 grid. It was hoped that a 'score to beat' index shown to the right of the screen, kept participants motivated in engaging in this NDRT.

Each of the two ~12-minute experimental drives contained 6 automation failures. This allowed 12 failures to be experienced by each driver in total. The timing of the failures within each drive was randomised to avoid drivers anticipating them, but they typically occurred within ~1-3 minutes of each other. To establish if road curvature influenced detection of failures, three failures were presented on Straight sections of the road, and three when the vehicle was on a Curved section of road. The order of Straight and Curved road segments was counterbalanced between conditions.

Participants engaged automation by pressing a button on the steering wheel. When automation was engaged, drivers were required to take their hands off the steering wheel and foot off the pedals, allowing the vehicle to drive at 70 mph, tracking the centre of the road, which contained a mixture of 252 m Straight and Curved sections. The Curved road sections had a radius of 1000 m. In addition to recording response to the Arrows task and number of VMS words cited, participants' response time to automation failures and control of the vehicle after failure was recorded.

The study took approximately 90 minutes to complete, and drivers were paid £20 for their time. Following the two drives, all participants completed an 8-part questionnaire, including the Driver

Behaviour Questionnaire, a record of their experience with in-vehicle technology, and their perceptions of the different aspects of the experiment. However, these are not reported here.

2.4. Automation Failure and Human-Machine Interface (HMI)

Unlike previous studies in this context, drivers were not provided with a take-over request (ToR) or any overt HMI-based notification of failure/limitation of the automated system. However, at the point of automation failure, the HMI indicated that the automated driving system was no longer in the Automation Engaged state, as shown in Figure 2.

The HMI, which was developed as part of the EC-funded AdaptIVe project, and based on a concept designed by CRF (Fiat; as first shown in Gonçalves et al., under review), was used to inform drivers about the engagement of the automated system (Figure 2), and appeared on the vehicle dash. The HMI could be used by drivers to check if the automated driving system was off.

Figure 2 - In-vehicle HMI, illustrating automation states.

Failures were introduced by a change in vehicle state from L2 automation (adaptive cruise control and lane keep assist) to full manual control. The failures coincided with a very subtle lateral movement of the vehicle, achieved by offsetting the steering wheel angle by 0.2 degrees at the point of automation failure. The salience of this cue was piloted to ensure that it was hard to detect, especially if drivers were looking away from the road scene.

Drivers were provided with the following written instructions regarding the automation failures:

> *"At designated points in the drive, and with no other warning, the automated system will fail. You do not need to 'look out' for these, as your main goal is to beat your score in the Arrows task. However, every time you notice such a failure, which is accompanied by a change in speed or position of the vehicle, you should try to resume control of the vehicle as soon as you can,*

position it in the correct lane and re-engage the automated system, continuing with the Arrows and VMS tasks, as required."

Some effort was invested in triggering of the automation failure at the correct road position, to ensure its presence on the Straight and Curved sections, while also preventing any link between VMS presence and automation failure.

2.5. Analysis

2.5.1. Vehicle-based driving performance

To establish if the proprioceptive feedback accompanying automation failure was noticeable by drivers, take-over time was calculated from the point at which the automation failed to the point at which drivers resumed manual control. Take-over time was determined on a case-by-case basis by the analyst, by inspecting drivers' response profiles in terms of steering wheel input, lateral accelerations, and vehicle lateral position. The method used here was originally developed by Louw, Markkula, Boer, Madigan, Carsten, & Merat (2017), and an example of such a classification method used in the current study is shown in Figure 3. The solid red line indicates the steering wheel signal, while the vertical dotted red line indicates the point at which the automation failure occurred. Looking to the left of the automation failure line, it is clear that the steering wheel input does not change. While there is a small change immediately after automation failure, there is no observable change to the vehicle's lateral position and lateral acceleration. Therefore, this was not deemed indicative of drivers taking over control. However, approximately 50 frames after the automation failure, there was a salient steering input that was accompanied by changes in vehicle offset and lateral acceleration. Therefore, take-over time was deemed the start of this steering input. This method was applied to each of the 12 take-over scenarios for the 30 drivers, resulting in 360 cases.

In addition to take-over time, to establish whether lane curvature, the presence of NDRT, and the VMS affected drivers' vehicle handling after failure, we also calculated the number of lane excursions after automation failure. A Lane Excursion is defined as a point when the centre of the driver's vehicle moved outside the intended driving lane into one of the adjacent lanes (Rumschlag et al., 2015). This was calculated from the point of automation failure until automation re-engagement.

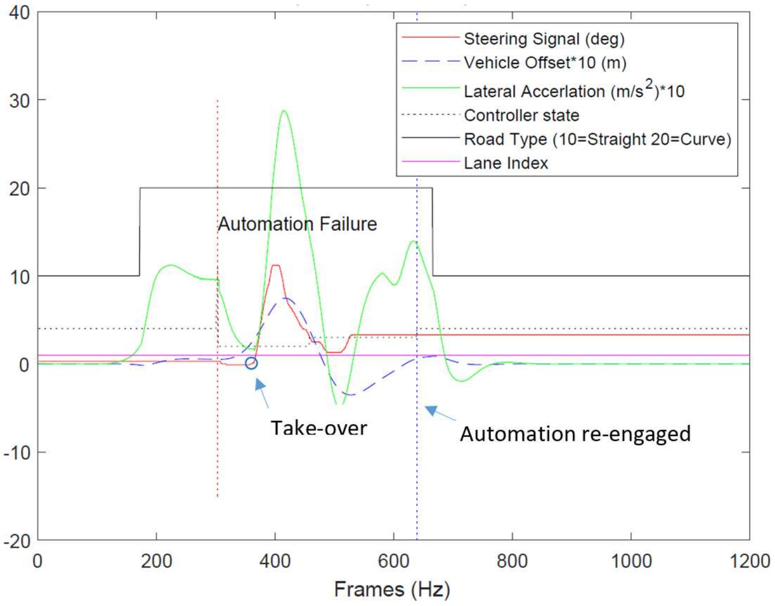

Figure 3 - Example plot from the analysis of a steering signal, to determine take-over time. Lateral acceleration and vehicle offset values are multiplied by 10 for illustration purposes.

2.5.2. Drivers' glance behaviour

To study the allocation of drivers' visual attention during automation and after automation failure, driver eye movements were analysed using a time-series representation of raw eye gaze, and the percentage and duration of eye fixations for different Areas of Interest (AoIs). Classification of gaze AoI was calculated based on valid intersections of gaze vectors with the predefined regions within the 3D world model. Subsequently, individual gaze points within the AoIs were grouped as fixations if they occurred within 5° in a 250ms window. 6 AoIs were identified: The road ahead region encompassed the forward roadway, the rear-view mirror and the VMS, the centre console region also encompassed the Arrows task, and the HMI region also encompassed the steering wheel and driver lap. Other regions included the left and right side mirrors. Additional classifications included off-road glances, which are those that are tracked but do not intersect with the other predefined regions, and unknown glances, which are those that are not tracked.

2.6. Statistical analyses

Data were analysed with SPSS V.24 (IBM, Armonk, New York, USA). Performance measures were analysed using a 2 (NDRT: VMS Only, VMS + Arrows) X 2 (Road Curvature: Curve, Straight) X 3 (Failure Order: 1st, 2nd, 3rd) analysis of variance (ANOVA), with both as within-participant factors. Fixation durations were compared using paired samples t-tests, while gaze transitions were analysed using a 2 (NDRT: VMS Only, VMS + Arrows) X 15 (Time Interval: 15 two second intervals around automation failure) analysis of variance (ANOVA), with both as within-participant factors.

An α-value of 0.05 was used as the criterion for statistical significance, and partial eta-squared was computed as an effect size statistic. Least significant difference pairwise comparisons (α=0.05) were used to determine the difference between levels of NDRT, Road Curvature, and Failure Order. In total, based on 30 participants driving the two roads, and six automation failures in each experimental drive (VMS Only and VMS + Arrows), there were 360 transition cases considered for the analysis.

3. Results and Discussion

All drivers resumed control from automation following all silent failures (n=360). To establish whether NDRT, Road Curvature, or Failure Order affected drivers' take-over responses, the results for each factor are compared and presented in three sections, 1) take-over time, 2) lane excursions, and 3) eye-movements.

3.1. Take-over Time (ToT)

There was a main effect of NDRT for Take-over Time (ToT), where drivers took significantly longer to resume manual control in the VMS + Arrows condition (M=1.49 s; SE=.08 s) compared to the VMS Only condition (M=2.18 s, SE=.22 s; F(1,28)=10.934, p=.003, η_p^2=.281; see Figure 4A). Therefore, engaging in the Arrows task seems to have delayed drivers' take-over response time, likely because they were not looking at the road and, therefore, did not notice the failure. This is in line with previous findings. For example, in their meta-analysis of determinants of ToTs in 93 studies, Zhang et al. (2018) report that drivers engaged in a visual-motor NDRT took longer (M=3.12 s; n=218) to take-over compared to when there was no NDRT (M=2.54 s, n=154). However, while the patterns are the same, ToTs in this study are substantially shorter, possibly because drivers were anticipating the automation failures. Moreover, the ToTs reported here probably over-estimate what drivers are capable of in real-world situations, especially if they are more intensely engaged in NDRTs, and do not expect, or are notified of, automation malfunctions.

There was also a main effect of Road Curvature, where drivers took significantly longer to respond to failures that occurred on Straight road segments (M=2.41 s; SE=.19 s) compared to those that occurred on Curved road segments (M=1.52 s, SE=.12 s; F(1,28)=9.222, p=.005, η_p^2=.248) - see Figure 4B). This was likely because, following an automation failure on a Curve, the vehicle did not maintain its previous heading angle, which is relative to the road angle. This provided drivers with a visual cue that was in addition to the proprioceptive cue, indicating the change in vehicle state because of the automation failure. However, if failures occurred on a Straight road segment, and the vehicle continued travelling on that Straight segment the apparent discrepancy between the vehicle's heading angle and the road angle was not as obvious to drivers, perhaps reducing their ability to notice the change in vehicle state, which accompanied the automation failure. This hypothesis is supported by a greater variability of take-over times in Straights than on Curves.

There was a main effect of Failure Order on ToT (F(1.656,46.357)=8.256, p =.002, η_p^2=.228). Drivers took significantly longer to respond in Failure 1 (M=2.21; SE=.16) compared to Failure 2 (M=1.71; SE=.16; p=.016) and Failure 3 (M=1.57; SE=.14; p<.001; Figure 4C), indicating that drivers learned to anticipate the automation failures. However, there were no two or three-way interactions, suggesting that drivers' take-over times did not adapt differently for Straight or Curves.

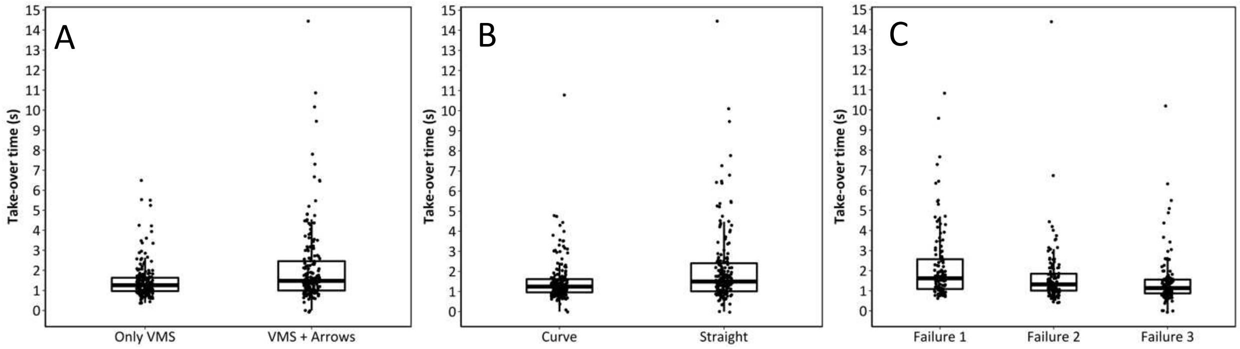

Figure 4 - Time from automation failure until the resumption of manual control for (A) the VMS Only and VMS + Arrows conditions, (B) when automation failed on Curved and Straight road segments, and (C) for each of the three failures.

3.2. Lane Excursions

In total, there were 50 Lane Excursions out of 356 valid cases (4 cases were excluded due to missing data). 70% (N=35/50) of these occurred during the VMS + Arrows condition, while significantly fewer (30%, N=15/50) occurred during the VMS Only condition ($\chi^2(1)$=9.8387, p=.001). Therefore, engagement in the Arrows task seems to have affected drivers' safety, as evidenced by their excursion to the adjacent lane in addition to longer ToTs.

Concerning Road Curvature, 64% (N=32/50) all Lane Excursions occurred when automation failed on Curved road sections, despite the shorter average take-over times in this condition (Figure

4B). This is likely due to the greater effort required to correct vehicle heading on a Curve road section after automation failure, compared to maintaining it on a Straight road section.

3.3. Analysis of Drivers' Eye-tracking

Exploratory analysis of drivers' eye-movements showed no differences with respect to road curvature or failure position. However, there were large differences between the VMS + Arrows condition and VMS Only condition, which possibly masked differences that may have existed between failures or between road curvatures. Therefore, in this section, we only compare the VMS + Arrows condition to the VMS Only condition. Results are presented in three sections, 1) before the automation failure, 2) after the automation failure, and 3) gaze transitions between AoIs over the full period.

To establish how drivers are dividing their attention in the moments around automation failure, Figure 5 shows the proportion of raw eye-gaze to the different AoIs, in the 10 s before and 20 s after the automation failure, for the VMS + Arrows and VMS Only conditions. To establish when, over time, drivers focus most on a particular AoI, Figure 6 shows the cumulative distribution of the total number of glances over the full time, for the most relevant AoIs, for each condition. These figures illustrate changes to how drivers are prioritising glances towards particular AoIs as the scenario unfolds.

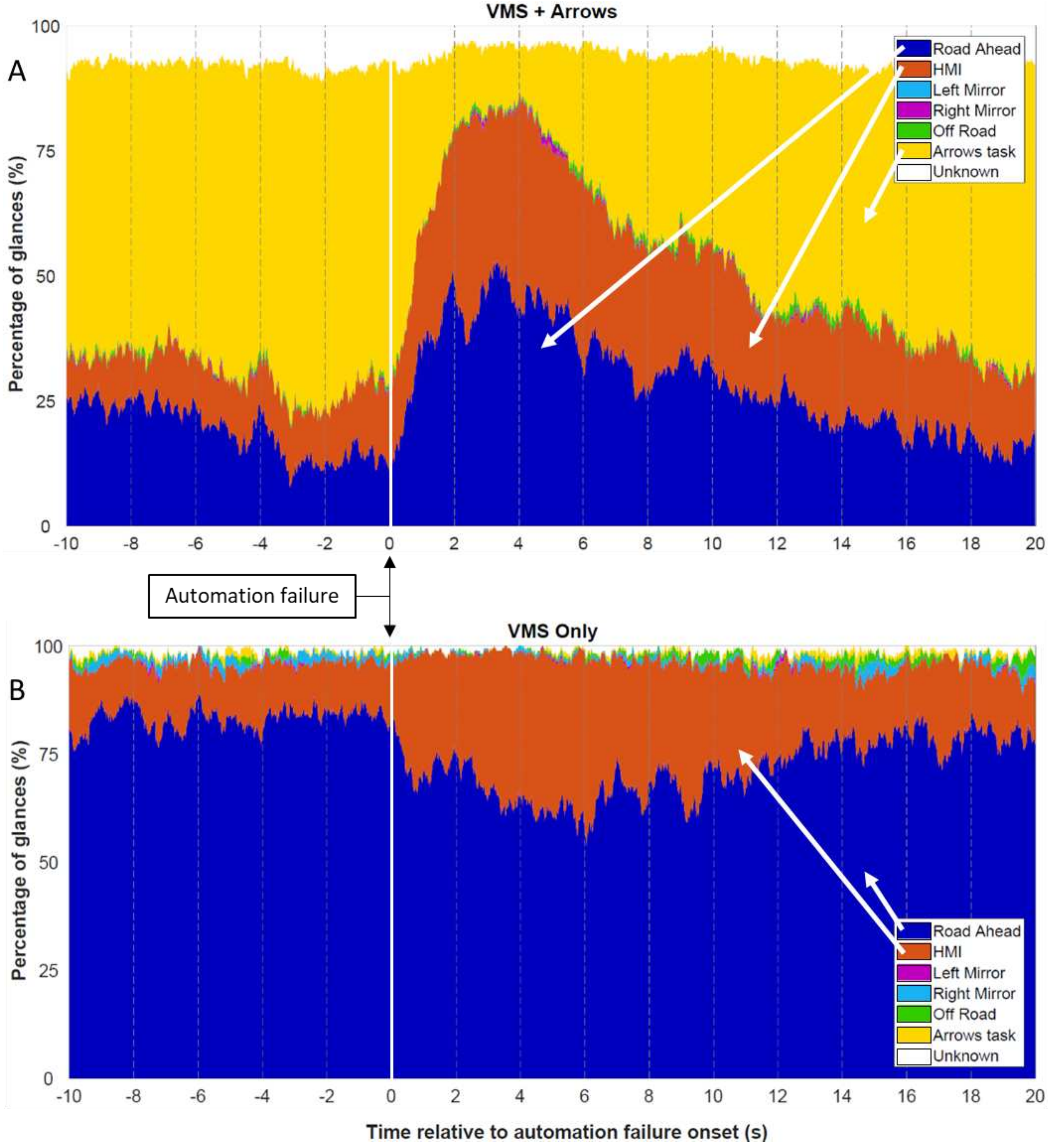

Figure 5 - Proportion of eye-gaze over time by Area of Interest for all drivers and all failures, in the (A) VMS + Arrows conditions and (B) VMS Only condition.

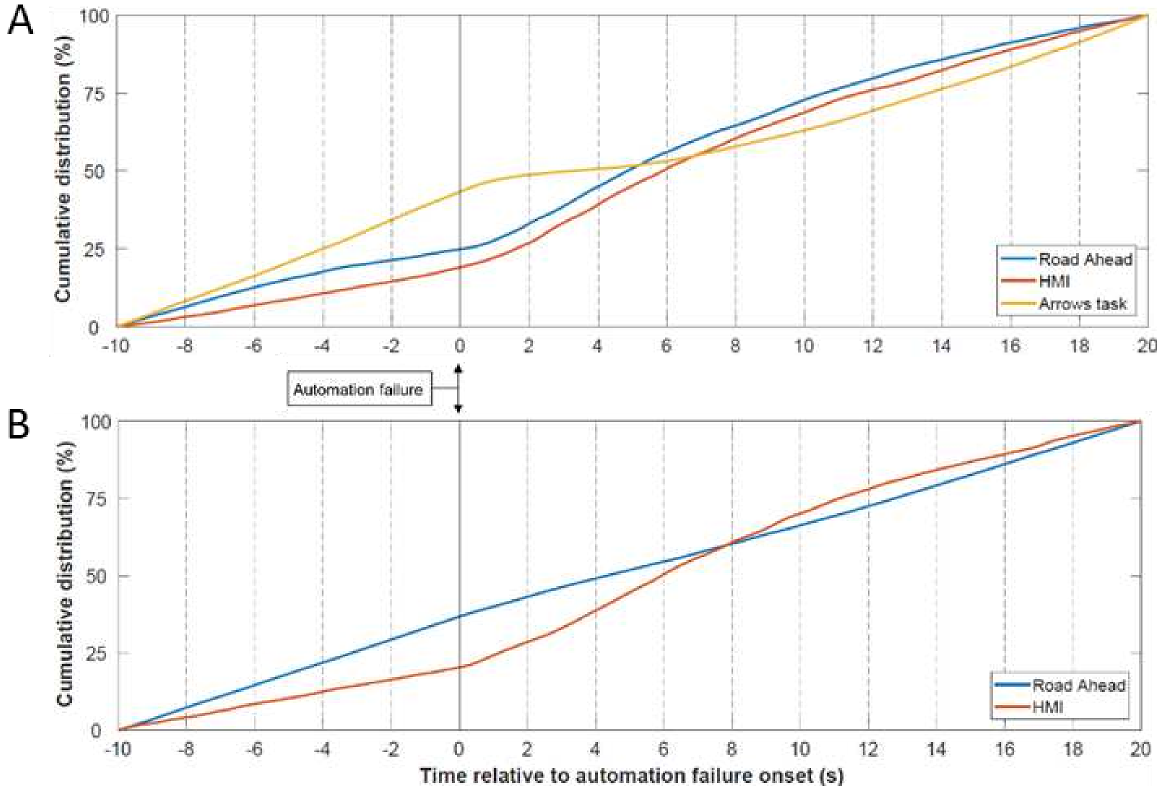

Figure 6 - Cumulative distribution of eye-gaze over time by Area of Interest for all drivers and all failures, in the (A) VMS + Arrows condition and (B) VMS Only condition.

To provide a more detailed insight into drivers' visual attention, we also calculated the location of drivers' fixations for the 10 s before the automation failure (Figure 7) and for the 2 s after automation failure. In addition, we also calculated the duration of fixations in the 2 s after automation failure.

While the above patterns show how visual attention is divided between the AoIs around the automation failure, and when, over time, drivers focused on each AoI most, it does not show the glance strategies underlying that division of attention. That is, whether drivers made more frequent glances or less frequent glances between the AoIs and whether this changes as a result of NDRT engagement or the automation failure. Therefore, we used raw gaze data, to calculate the average number of gaze transitions between the different AoIs in 2 s intervals from 10 s before until 20 s after the automation failure (Figure 9).

Before Automation Failure

The raw-gaze data in Figure 5A shows that in the 10 s before the automation failure, where the VMS was clearly visible, drivers in the VMS + Arrows condition focused predominantly on the Arrows task, with Figure 6A confirming that 50% of fixations were located towards the centre console AoI. The road ahead and HMI regions captured the second and third largest proportion of glances, with 34% and 9% of fixations, respectively. There were very few glances towards the mirrors during this period (Figure 7). This demonstrates that while the majority of drivers' attention was directed towards the Arrows task, they were still maintaining some attention to the road ahead and the HMI. However, as with ToTs, this level of glances towards the road ahead and HMI may not necessarily reflect those in real-world settings, because, as discussed in the previous section, drivers were engaging in the VMS task and possibly anticipating the automation failures.

Figure 5B shows that for the VMS Only condition, glances were mostly towards the road ahead during the 10 s prior to automation failure, with Figure 7 showing that approximately 75% of fixations were observed in this region. The second highest proportion of glances was for the HMI, which accounted for 19% of fixations, and the remaining fixations were towards the side mirrors (Figure 7).

Therefore, the addition of the Arrows task affected drivers' attention distribution, with fewer glances towards the side mirrors during this drive. The implications of these results on road safety are evident since a failure to attend to adjacent lanes on the run-up to an automation failure can be detrimental if drivers fail to maintain their vehicle in its lane. This reduction in visual attention to mirrors by NDRT and its effect on road safety has also been shown in manual driving studies (e.g. Recarte and Nunes, 2000).

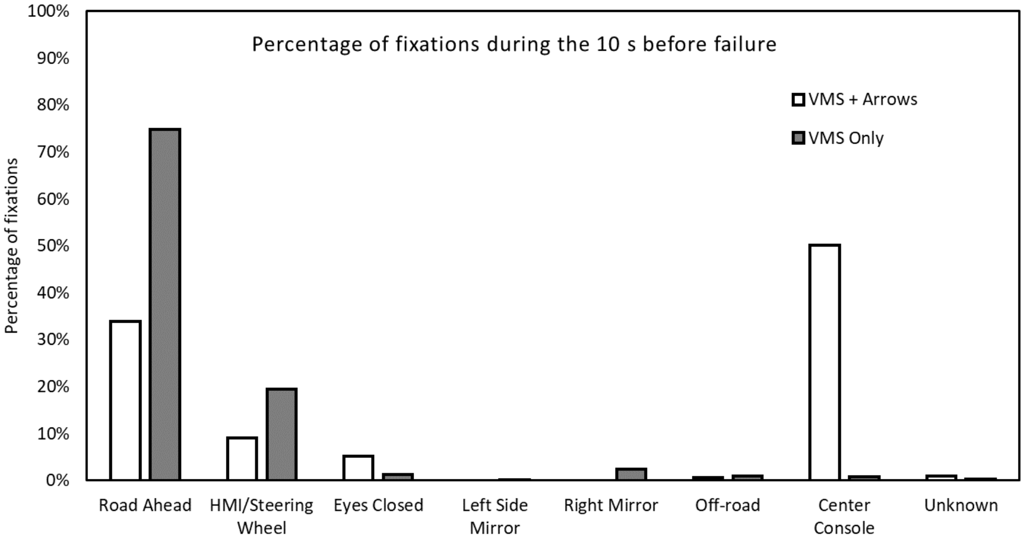

Figure 7 - Percentage of fixations for each Area of Interest in the 10 s before the automation failure, for the VMS Only and VMS + Arrows Task conditions, across all failures.

After Automation Failure

During the VMS + Arrows drive, there was a significant redirection of drivers' visual attention from the Arrows task towards the road ahead and (to a lesser extent) the HMI (Figure 5A). However, there were almost no glances towards the side mirrors during this period, suggesting that drivers had little spare capacity to scan the road environment, as they assessed the automation status and started the process of resuming manual control. Even as drivers returned to the Arrows task from about 5 s after the failure, the cumulative distribution in Figure 6A indicates that the proportion of attention sharing between the road ahead and HMI held until the end of the period under observation.

In addition, for the VMS Only condition, there was an increase in glances towards the HMI in the first 2 s after automation failure (Figure 5A). This is also illustrated by the noticeable increase in the cumulative distribution of HMI (Figure 6B), where this shift in attention resulted in a smaller proportion of glances towards the mirrors and the road ahead, though the road ahead still captured the majority of drivers' glances during this period. This figure also shows that drivers' focus towards the HMI continued until approximately 6 s after the automation failure, with a gradual increase in glances towards the road ahead and side mirrors. This is likely because, after this point, drivers felt either the HMI offered no new information or they re-engaged automation and there was no longer a need to focus so heavily on the HMI.

To better understand how drivers divided their attention during the first 2 s after automation failure, where there was a major redirection of their visual attention, we calculated the percentage and duration of fixations for each AoI, for the VMS Only and VMS + Arrows Task conditions, across all failures (Figure 8). Results showed that, in the absence of the Arrows task, drivers tended to fixate more towards the HMI/Steering wheel in the first 2 s after automation failure, as compared to the preceding 10s (from 19% to 33%), at the cost of attention towards the road ahead (from 75% to 64%; Figure 8A). Therefore, drivers were monitoring the status of the automated driving system more in the moments after the automation failure, which suggests that the correct and easily understood/visible message provided by HMI is important for such scenarios. To ensure driver visual attention towards the HMI does not detract from important road-based events, more consideration must be devoted to the type and position of such messages.

For the VMS + Arrows condition, the percentage of fixations towards the road ahead was found to increase from 34% to 54% in the 2 s after failure, when compared to the 10 s period before the failure (Figure 8A). Similarly, there was an increase in fixations towards the HMI/Steering wheel, from 9% to 25%. However, while there was a reduction in fixations towards the Arrows task, the significantly longer take-over times shows that the Arrows task clearly prevented drivers from acknowledging or attending to the automation failure.

This is further illustrated by the duration of drivers' fixations, shown in Figure 8B. We calculated a paired samples t-test to compare the duration of drivers' fixations that occurred in the 2 s after the automation failure, in the two conditions. Drivers' fixation duration towards the road ahead was significantly longer in the VMS only condition, compared to the VMS + Arrows drive (M=1.69 s, SE=.23, vs. M=1.01 s, SE=.10; t(28)= -8.2569, p<.001), suggesting that Arrows engagement resulted in fewer and shorter fixations to the road ahead in the moments after the automation failure. However, there was no difference in the duration of fixations to the HMI.

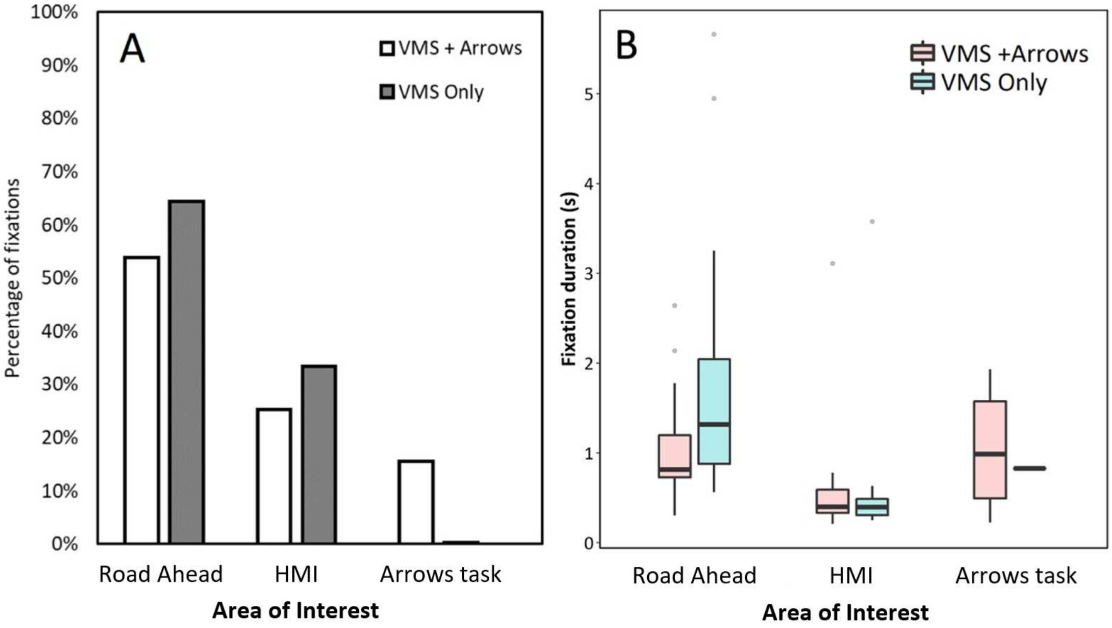

Figure 8 - Percentage of fixations (A) and the average duration of fixations (B) for three of the main Areas of Interest in the first 2 s after the automation failure, for the VMS Only and VMS + Arrows Task conditions, across all failures.

Gaze transitions between AoIs

There was a main effect of NDRT on the number of gaze transitions, where there were more gaze transitions between AoIs in the VMS + Arrows (M = 2.85, SE = 1.55) condition compared to the VMS Only condition (M=2.26, SE=.07; F(1,173)=14.757, p<.001, η_p^2=.079), with Figure 9 showing that the difference between the two conditions similar throughout. In addition, there was much greater variability between drivers in the VMS + Arrows condition, compared to the VMS Only condition, suggesting different coping strategies by drivers. There was also a main effect of Time Interval ($F(14,2422)=14.757$, $p<.001$, $\eta_p^2=.079$), but no interaction.

It is perhaps unsurprising that there are more gaze transitions in the VMS + Arrows condition, as the Arrows task gives drivers an additional point of visual attention. However, perhaps unexpected is that the difference in the average number of transitions between the two conditions remained similar. This suggests that whether or not drivers engaged in an NDRT did not affect how much more (or less) they increased the rate at which they moved their eyes between the different AoIs after the automation failure.

However, the data do indicate that NDRT engagement results in larger individual differences in gaze patterns, where some drivers may prefer to make more frequent gazes between AoIs, while others prefer to make less frequent glances. In other words, as drivers' attention is divided between more information sources, the less predictable is their gaze strategy, which suggests that not all drivers can be relied upon to monitor and detect deficiencies in system performance. Moreover, for both drives, the number of drivers' gaze transitions increased similarly after the automation failure. However, where the average number of gaze transitions in the VMS Only condition began to decline in the 6-8 s interval, this occurred 2 s later in the VMS + Arrows condition. The delay is explained by the longer take-over times in the VMS + Arrows condition (Figure 4A), compounded by the fact that some drivers were clearly still engaging in the Arrows task while dealing with the automation failure (Figure 5A).

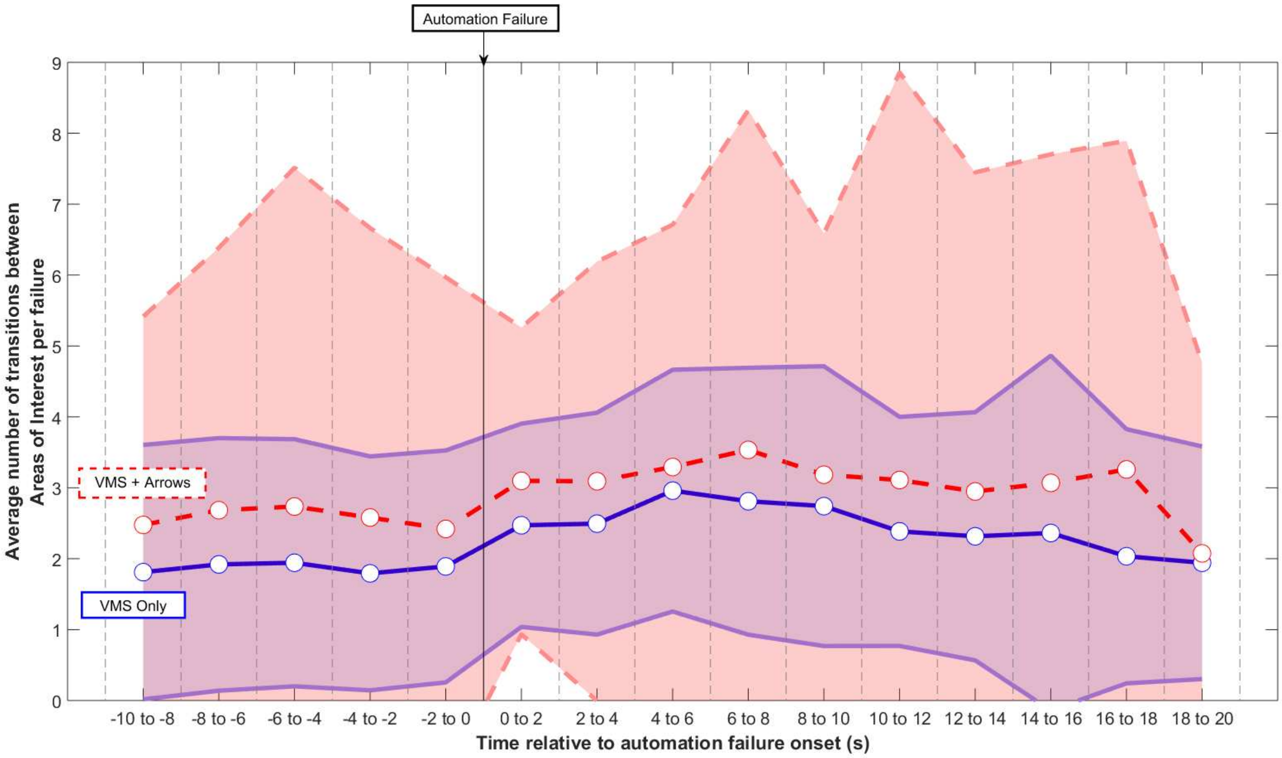

Figure 9 - The average number of gaze transitions between the different AoIs, per driver, per failure, for the VMS Only and VMS + Arrows condition. Shaded areas represent the standard deviation.

4. Summary and Conclusions

The objective of this experiment was to investigate the effects of engaging in an NDRT on driver glance strategies before and after silent automation failures. The results address several aspects of this objective.

First, the data demonstrate that drivers can detect and respond to a subtle failure of the automated driving system, cued by changes in kinematic feedback from the vehicle, and in the absence of any obvious ToRs, particularly if drivers are not engaged in an NDRT. Compared with only looking at the road ahead during automation, engaging in an additional NDRT during automation resulted in significantly more lane excursions and longer take-over times. While the take-over times presented here were surprisingly short, given there were no ToRs, there are some reasons why these may still be an over-estimation of drivers' behaviour in real-world settings. In real-world settings, drivers may not be anticipating failures, and may even delay their intervention to assess whether the automation is indeed malfunctioning. In addition, in real-world settings drivers may be more disengaged from the road environment when interacting with NDRTs, especially as there will not necessarily be the motivation to direct attention towards the road, as encouraged here by the VMS task. This may result in slower responses and less frequently detected automation failures. A counterpoint, however, is that the actual safety threat posed in real-world scenarios may induce a more engaged and responsive driver than does an artificial threat in simulator experiments, and this is also likely to be affected by drivers' accepted level of risk. Moreover, increasingly, vehicle manufacturers are incorporating stricter hands-on requirements into the interaction designs of their L2 systems. This forces more regular engagement in the driving task than what has been simulated here, which could improve the chances that drivers will detect subtle perturbations in system functionality.

Second, while failures on Curved road segments resulted in shorter take-over times, compared to in Straights, there were also more lane excursions, demonstrating the importance of accounting for context when assessing take-over performance, and not only relying on one surrogate measure of take-over quality.

Third, the data show a clear difference in glance behaviour between the VMS Only and VMS + Arrows conditions during automation, where the majority of drivers' glances were towards the road ahead and Arrows task, respectively. In both conditions, drivers maintained some attention towards the HMI. However, the spare capacity afforded to drivers in the VMS Only condition meant that they paid more attention to the HMI and could make regular glances towards the mirrors, where this was not the case in the VMS + Arrows condition.

Finally, there were differences between the conditions in terms of how much drivers had to change the focus (and distribution) of their visual attention following a silent automation failure, and, importantly, how long it took for this change to happen.

There were more lane excursions in the Arrows condition, especially in the 10 s before automation failure, where the majority of drivers' glances were on that task, which may have contributed to lane excursions. However, there were still 15 lane excursions in the in VMS Only condition, where the majority of drivers' visual attention was towards the road ahead in the lead up to the automation failure. While not a like-for-like comparison, our results echo those of an on-road test-track study by Victor et al. (2018), where even with their eyes on the conflict object, over a quarter of drivers crashed with a stationary car or garbage bag. This was despite drivers receiving explicit and detailed instructions on system-specific limitations in risk scenarios. This highlights the importance of providing accurate mental models of the system limitations to drivers, supported by practical, and not only theoretical, training.

Although this study has demonstrated how drivers' responses and glance patterns around silent automation failures is affected by engaging in NDRTs, several questions remain unanswered. This study provided drivers with a very limited exposure to the situations that might give rise to automation failures or take-over requests. It did not, however, systematically explore how variations in situation criticality might interfere with glance strategies. Previous studies have shown a link between situation criticality (viz., varying take-over time budget) and take-over performance. However, a more comprehensive set of studies needs to be undertaken to establish how different situations elicit different glance patterns and whether these can be used to predict whether drivers will safely take over control. Here, the challenge is not only extracting the appropriate glance metrics but also defining an appropriate measure of take-over performance, one that considers in concert the context and quality of driver response (Louw, Markkula, Boer, Madigan, Carsten, & Merat, 2017; Saïd & Chauvin, 2017). This will help inform more driver- and situation-specific support strategies during take-over situations, in terms of what and how information is presented via the HMI or the level of vehicle control support drivers receive.

Finally, further research is required to understand how best to encourage drivers to maintain consistent and meaningful attention towards the forward roadway, because, as shown previously, drivers tend to disengage more from the driving task with higher levels of automation (Carsten et al., 2012). One strategy for increasing drivers' attention towards the forward roadway, and their ability to detect deficiencies in system performance, is to provide drivers with real-time feedback regarding their road-monitoring level (e.g. Cohen-Lazry, Borowsky, & Oron-Gilad, 2017), for example, in the preceding 10 seconds. This could be further motivating via gamification, by providing a monitoring score against a pre-determined goal or threshold, an approach used successfully to encourage eco-driving behaviour (Hibberd, Jamson, & Jamson, 2015; Magana & Munoz-Organero, 2015).

5. Acknowledgements

This research was funded by Seeing Machines Inc., Canberra, Australia.

6. References

Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics*, 68, 138-145.

Benmimoun, M., Pütz, A., Zlocki, A., & Eckstein, L. (2013). eurofot: Field operational test and impact assessment of advanced driver assistance systems: Final results. In *Proceedings of the FISITA 2012 World Automotive Congress* (pp. 537-547). Springer, Berlin, Heidelberg.

Cadillac (2018). *Introducing Cadillac Super Cruise*. Retrieved from: https://www.cadillac.com/world-of-cadillac/innovation/super-cruise

Carsten, O., Lai, F. C., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control task substitution in semiautomated driving: Does it matter what aspects are automated? *Human factors*, 54(5), 747-761.

Cohen-Lazry, G., Borowsky, A., & Oron-Gilad, T. (2017, September). The effects of continuous driving-related feedback on drivers' response to automation failures. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 1980-1984). Sage CA: Los Angeles, CA: SAGE Publications.

Endsley, M. R. (2017). Autonomous driving systems: A preliminary naturalistic study of the Tesla model S. *Journal of Cognitive Engineering and Decision Making*, 11(3), 225-238.

Eriksson, A., & Stanton, N. A. (2017). Takeover time in highly automated vehicles: noncritical transitions to and from manual control. *Human factors*, 59(4), 689-705.

Feldhütter, A., Gold, C., Schneider, S., & Bengler, K. (2017). How the duration of automated driving influences take-over performance and gaze behavior. In *Advances in Ergonomic Design of Systems, Products and Processes* (pp. 309-318). Springer, Berlin, Heidelberg.

Flemisch, F., Altendorf, E., Canpolat, Y., Weßel, G., Baltzer, M., Lopez, D., ... & Schutte, P. (2017). Uncanny and unsafe valley of assistance and automation: First sketch and application to vehicle automation. In *Advances in Ergonomic Design of Systems, Products and Processes* (pp. 319-334). Springer, Berlin, Heidelberg.

Gonçalves R.C., Louw, T., Madigan, R., Quaresma, M. & Merat, N. (Under review). Drivers' eye-gaze pattern during lane-changing in highly-automated driving.

Hibberd, D. L., Jamson, A. H., & Jamson, S. L. (2015). The design of an in-vehicle assistance system to support eco-driving. *Transportation Research Part C: Emerging Technologies*, *58*, 732-748.

Huisingh, C., Griffin, R., & McGwin Jr, G. (2015). The prevalence of distraction among passenger vehicle drivers: a roadside observational approach. *Traffic injury prevention*, 16(2), 140-146.

Jamson, A. H., & Merat, N. (2005). Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(2), 79-96.

Lambert, F. (2018). *Tesla Autopilot confuses markings toward barrier in recreation of fatal Model X crash at exact same location*. Electrek. Retrieved from: https://electrek.co/2018/04/03/tesla-autopilot-crash-barrier-markings-fatal-model-x-accident/

Lee, J., Moeckli, J., Brown, T., Roberts, S., Victor, T., Marshall, D., ... & Nadler, E. (2013, May). Detection of driver distraction using vision-based algorithms. In *Proc. 23rd Enhanced Safety Veh. Conf*. Seoul, South Kora, 2013.

Li, S., Blythe, P., Guo, W., & Namdeo, A. (2018). Investigation of older driver's takeover performance in highly automated vehicles in adverse weather conditions. *IET Intelligent Transport Systems*, 12(9), 1157-1165.

Llaneras, R., Salinger, J., & Green, C. (2013). Human Factors Issues Associated With Limited Ability Autonomous Driving Systems: Drivers’ Allocation Of Visual Attention To The Forward Roadway. In *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (pp.92-98). Iowa City, Iowa: The University of Iowa, Public Policy Center.

Louw, T., Kountouriotis, G., Carsten, O., & Merat, N. (2015). Driver inattention during vehicle automation: how does driver engagement affect resumption of control? In *4th International Conference on Driver Distraction and Inattention*, 9-11 November 2015, Sydney, Australia.

Louw, T., Madigan, R., Carsten, O., & Merat, N. (2017). Were they in the loop during automated driving? Links between visual attention and crash potential. *Injury prevention*, 23, 281-286.

Louw, T., Markkula, G., Boer, E., Madigan, R., Carsten, O., & Merat, N. (2017). Coming back into the loop: Drivers’ perceptual-motor performance in critical events after automated driving. *Accident Analysis & Prevention*, 108, 9-18.

Louw, T., Merat, N., & Jamson, H. (2015). Engaging with highly automated driving: To be or not to be in the loop? In *Proceedings of 8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Salt Lake City, Utah, USA.

Madigan, R., Louw, T., & Merat, N. (2018). The effect of varying levels of vehicle automation on drivers' lane changing behaviour. *PloS one*, 13(2), e0192190.

Magana, V. C., & Munoz-Organero, M. (2015). GAFU: using a gamification tool to save fuel. *IEEE Intelligent Transportation Systems Magazine*, *7*(2), 58-70.

Marshall, A., & Davies, A. (2018, May 24). *Uber's self-driving car saw the woman it killed, report says.* Wired. Retrieved from https://wired.com/story/uber-self-driving-crash-arizona-ntsb-report/

Melcher, V., Rauh, S., Diederichs, F., Widlroither, H., & Bauer, W. (2015). Take-over requests for automated driving. *Procedia Manufacturing*, 3, 2867-2873.

Merat, N., Seppelt, B., Louw, T., Engström, J., Lee, J. D., Johansson, E., ... & McGehee, D. (2018). The "out-of-the-loop" concept in automated driving: Proposed definition, measures and implications. *Cognition, Technology & Work*, 1-12.

Naujoks, F., Mai, C., & Neukum, A. (2014). The effect of urgency of take-over requests during highly automated driving under distraction conditions. *Advances in Human Aspects of Transportation*, 7(Part I), 431.

Naujoks, F., Purucker, C., & Neukum, A. (2016). Secondary task engagement and vehicle automation–Comparing the effects of different automation levels in an on-road experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, 38, 67-82.

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of experimental psychology: Applied*, *6*(1), 31.

Rumschlag, G., Palumbo, T., Martin, A., Head, D., George, R., & Commissaris, R. L. (2015). The effects of texting on driving performance in a driving simulator: The influence of driver age. *Accident Analysis & Prevention*, 74, 145-149.

Russell, H. E., Harbott, L. K., Nisky, I., Pan, S., Okamura, A. M., & Gerdes, J. C. (2016). Motor learning affects car-to-driver handover in automated vehicles. Science Robotics, 1(1).

SAE (2016). *Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems (J3016).* Technical Report Technical report, SAE International.

Saïd, F., & Chauvin, C. (2017, July). Automated vehicles: multivariate analysis of drivers' take-over behaviour. *In 13th ICNC-FSKD Conference 2017.*

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human factors*, 60(8), 1095-1116.

Volvo (2018). IntelliSafe. *A factsheet on Volvo Cars' safety technology for Model Year 2017*. Retrieved from: https://assets.volvocars.com/ru/~/media/russia/about/our-stories/news-img/2016/april/180996_IntelliSafe_factsheet_for_the_MY2017_Volvo_line_up.pdf

Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis & Prevention*, 92, 230-239.

Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (Under review). Determinants of take-over time from automated driving: A meta-analysis of 129 studies.