Department for Transport

Regaining Situational Awareness as a User in Charge: Responding to transition demands in automated vehicles

January 2025

Department for Transport
Great Minster House
33 Horseferry Road
London
SW1P 4DR
enquiries@ccav.gov.uk

Control sheet

Project Manager	Dr Clare Mutzenich	
Quality Assurance	Hannah Hughes	10th September 2024
Version Control	Draft 1	31st July 2024
	Draft 2	27th August 2024
	Final report	11th September 2024
Date of Approval	Stephen Devlin	7th January 2025

This report has been prepared by Lacuna Agency, a strategic insights agency serving clients in public sector, automotive, luxury, and sport. We combine scientific techniques with evidence-based practices to design innovative research for hard-to-reach or difficult-to-study audiences. Our work also extends to the broader mobility and transport sectors, where we assist governments, manufacturers and network providers in planning for future needs.

Executive Summary

Background to the Project

In 2022, the Law Commission of England and Wales and the Scottish Law Commission (The Law Commissions) published a report suggesting a substantial overhaul of the legal framework governing automated vehicles. This joint report conducted a comprehensive review of the regulatory structure for automated vehicles on public roads and highways and introduced a novel legal entity known as the User-in-Charge (UiC), an individual situated within the vehicle and capable of operating the driving controls while a self-driving feature is engaged but not a driver (Law Commissions, 2022a). The UiC does not have to monitor the road while the self-driving system is active.

This project explores the transition from self-driving technology to human-operated driving, particularly focusing on the role of the UiC, who must be ready to take control when the automated system issues a takeover request. A takeover request occurs when the automated driving system encounters emergencies or conditions outside its programming. The UiC receives the takeover request via visual, auditory, or haptic cues.

The first approved system under these guidelines is the Automated Lane Keeping System (ALKS), which requires the UiC to take over within 10 seconds of a takeover request. As self-driving technologies become more prevalent, questions arise regarding the permissibility of engaging in Non-Driving Related Activities (NDRA) while using vehicles with activated self-driving features such as ALKS. There is potential for certain NDRAs to be allowed, provided they do not compromise the driver's ability to resume control of the vehicle safely when a takeover request is issued by the system.

Upon receiving a takeover request, the UiC must suddenly focus on building Situational Awareness (SA) from the surrounding road environment to enable a safe takeover. SA is critical for safe driving, including during the transition from automated to manual control. SA involves three levels: perception of the environment, comprehension of the situation, and projection of future events. The time required to gain sufficient SA after a takeover request is crucial, with studies showing response times ranging from 3 to 20 seconds. While simple tasks, like basic perceptual awareness, are processed quickly, more complex activities that involve higher levels of situational awareness—such as interpreting road signs or anticipating how current events will impact the road ahead—require more time.

Project Aims

This project focused on exploring the implications of NDRAs in vehicles with self-driving capabilities, when a UiC must respond to a transition demand. As the use of ALKS becomes more prevalent, it is essential to understand which NDRAs can be performed (within the 10 seconds mandated by ALKS regulations) without compromising the ability to safely resume control of the vehicle.

The project had the following objectives:

- Investigate which non-driving related activities (NDRAs), if any, can be safely performed during non-driving periods in cars with self-driving features but requiring a UiC.
- Establish mechanisms for measuring SA and determine appropriate thresholds to ensure safe takeover and resumption of manual driving.
- Understand potential variations in the impact of NDRAs across different scenarios to inform policy development.

To achieve these objectives, Lacuna Agency worked in partnership with University College London (UCL) and Loughborough University (LU), combining our expertise in Human Factors research and driving simulation. The project used a simulator-based approach to closely replicate real-life driving conditions. This collaboration, together with DfT and CCAV aimed to provide evidence-based insights that will inform future policy development and contribute to safer implementation of automated driving technologies.

Method

The research was conducted using high-fidelity driving simulators at two locations: University College London (UCL) and Loughborough University (LU), with 97 participants representing the general UK driving population. The study used a within-participants design with eight trials per participant, involving two motorway scenarios—roadworks and congestion—designed to simulate conditions that exceed the operational limits of the ALKS and trigger planned takeover requests due to speed changes:

1. **Roadworks Scenario**: The ego vehicle (which the participant controls in the driving simulator) drove at 68 mph in light traffic. After 2 to 4 minutes, a roadworks sign appeared, prompting a takeover request. Participants were expected to decelerate upon taking manual control in response to speed limit signs.
2. **Congestion Scenario**: The ego vehicle drove at 37 mph due to traffic congestion. After 2 to 4 minutes, the congestion cleared triggering a takeover request. The participants were expected to accelerate to match the surrounding traffic speed.

Participants were briefed on the study's aims and procedures and completed a pre-clinic questionnaire assessing their familiarity with technology and attitudes toward automated driving. After fitting the participants with eye-tracking glasses (and EEG caps at UCL), they engaged in a practice drive to familiarise themselves with the simulator controls and were assessed for simulator sickness.

Each trial involved an automated driving phase, during which participants were engaged in one of the following NDRAs or "No NDRA" conditions:

Mobile Phone Activities:

- Watching a Film: Participants selected a 5-minute YouTube video from a range of options (e.g., a TED talk, a nature documentary) to watch on a Google Pixel 6a smartphone, which was cradled on the dashboard.
- Playing Tetris: Participants played Tetris on a handheld Google Pixel 6a smartphone. The phone was placed on the passenger seat, and participants were instructed to pick it up and start playing once the trial began.

Non-Technological Activities:

- Reading a Magazine: Participants chose from a selection of magazines (e.g., BBC Top Gear, National Geographic) and were instructed to pick up the magazine from the passenger seat and start reading once the trial began.
- Completing a Wordsearch or Sudoku: Participants were given the choice between a word search puzzle book or a Sudoku book, along with a pen. The materials were placed on the passenger seat, and participants were instructed to start solving the puzzle once the trial began.

Motoric Activities:

- Drinking Water: Participants drank water from a disposable coffee cup with a lid, placed in the cup holder beside the driver. They were instructed to take frequent small sips throughout the drive.
- Simulated Eating of “Popcorn”: Participants mimicked the action of eating by transferring cotton balls from a packet into a cup holder attached to their chest. This simulated the action of eating popcorn, avoiding the mess and potential distractions of real food.

In each trial, participants received instructions on how to carry out the NDRA (or were informed there would be no NDRA). The system operated in self-driving mode, and participants were told to take manual control after a takeover request, "*as soon as you feel ready and safe to do so*."

Automated drives lasted 2-4 minutes, ending with a takeover request. Takeover requests were signalled by an auditory beep and a visual alert on the Human Machine Interface (HMI). Participants had 30 seconds to take control. At UCL, manual control was regained when the participant used the steering wheel or pedals, while at LU, participants were required to say, “Ready to drive” and a researcher gave manual control. Participants then drove manually for 30 seconds before completing a questionnaire on workload and self-perceived SA.

Data Collection and Analysis

A comprehensive set of measures were employed to assess SA following a takeover request during automated driving. The analysis was designed to investigate whether participants fully disengaged from NDRAs, their visual attention patterns post-takeover, particularly focusing on whether they looked at mirrors or other critical areas of the driving environment, and whether participants took appropriate behavioural actions, such as adjusting speed based on the scenario (which would indicate their comprehension of the reasons behind the takeover request). Additionally, the quality of the takeover was assessed by observing any signs of erratic steering, such as swerving or crashing, which could suggest either poor control or insufficient SA.

GoPro cameras captured in-cab behaviours, allowing for detailed analysis of how participants managed the transition from NDRA to driving. Participants’ interactions with the NDRA, their disengagement process, and their subsequent driving performance were all recorded. This comprehensive data collection approach enabled the analysis of how different NDRAs influenced the ability of participants to regain SA and safely resume control of the vehicle.

Key Findings

Some NDRAs were easier to disengage from than others. Activities like reading a magazine resulted in quicker and more consistent takeover times, as participants could easily disengage from them. However, mobile phone tasks had lower disengagement rates, with participants often continuing the activity even when initiating manual driving. This led to quicker takeover times but did not guarantee a safe transition, as participants may not have fully understood the driving actions required.

Ability to safely take over within regulated timeframes was variable. Although participants were not informed about the 10-second ALKS regulation, many attempted to take over quickly. However, some took longer than 10 seconds, either due to careful disengagement or slower responses. This variability raises concerns about the adequacy of the 10-second takeover period. Even simple tasks, such as putting a lid on a pen, can negatively affect a driver’s ability to resume control in a safe and timely manner.

Driving scenario complexity impacts takeover times. Takeover performance varied depending on the driving scenario. Roadworks scenarios, in particular, resulted in slower and more variable responses compared to congestion scenarios. Participants may have struggled to pick up on critical environmental cues, such as road signs or traffic changes. Certain NDRAs, like using a cradled mobile phone, eating popcorn, or doing a wordsearch, significantly delayed response times in roadworks but had a lesser effect in congestion. Although the difference between the two scenarios was minimal (less than half a second), at motorway speeds, even this delay could be dangerous, covering approximately 15.4 metres in just half a second.

Individual differences and environmental factors play a significant role. There was noticeable variability in takeover performance between participants and locations (UCL vs LU). UCL participants generally took longer to reach target speed, particularly in roadworks scenarios. Activities such as using a cradled mobile phone or completing a wordsearch caused more delays at UCL compared to LU. This difference could be attributed to variations in the simulator environments or methods of takeover between the two locations. However, no significant interaction effects were found between location and performance, indicating that while individual and environmental factors matter, they did not conclusively impact performance outcomes. Similarly, although no significant effects of specific NDRAs on lane deviations were found, the variability between participants highlights the importance of considering individual and environmental factors when assessing takeover performance.

Mirror checks were rarely used to build SA following a takeover request. Eye-tracking data showed that participants rarely used mirror checks to build situational awareness (SA) after receiving a takeover request. Instead, they primarily focused on the road and speedometer, with some looking to the HMI for information. This suggests that participants may have been unsure of where to find critical information. Enhancing the HMI to provide clearer, context-specific details about the takeover request could help participants build SA more quickly, leading to safer transitions from automated to manual driving.

Clearer guidance on what constitutes a safe and effective takeover is needed. Participants often struggled to properly disengage from NDRAs or did not recognise that continuing the activity compromised takeover safety. The vague instructions provided—"*take over as soon as you feel it is safe to do so*"—likely contributed to inconsistent interpretations of what constitutes a safe takeover. Clearer guidance for self-driving vehicles is needed, including explicit instructions on disengaging fully from tasks and building SA before resuming control. This would help ensure a safe and effective takeover process during transitions from automated to manual driving.

Conclusions

While some NDRAs may be safely performed during periods of self-driving in an automated vehicle, many can significantly impair SA and delay the transition to manual control, particularly in complex driving scenarios like roadworks. This project draws attention to the need for refined mechanisms to measure SA and establish appropriate thresholds for safe takeovers and the importance of providing clear and specific instructions to drivers in automated vehicles to ensure that they understand how to conduct a safe takeover. This includes not just taking control quickly but doing so in a manner that ensures they build sufficient SA to resume manual driving safely. The variability in participant responses and the influence of environmental factors suggest that further research is necessary to fully understand the nuances of NDRA impacts across different scenarios. This ongoing research will be crucial for developing informed policies and enhancing the safety of automated driving systems.

Contents

Glossary of terms

Key term	Definition
Automated Driving System (ADS)	Vehicle system using hardware and software to perform the entire dynamic driving activity. Can be considered self-driving. (BSI, 2020)
Automated Lane Keeping System (ALKS)	Hardware and software designed for low-speed applications which require driver activation and are responsible for keeping the vehicle within its lane at speeds of 60kph or lower on motorways. These systems control both lateral and longitudinal movements of the vehicle over prolonged durations, eliminating the need for additional input from the driver. Periods of self-driving but User-in-Charge must take back control when system issues a transition demand (United Nations, 2021).
Electroencephalogram (EEG)	A test that measures the electrical activity of the brain. This is done using small sensors called electrodes, which are placed on the scalp. These electrodes detect the tiny electrical signals produced by brain cells communicating with each other. The recorded brain wave patterns are then analysed (Yang et al., 2018).
NASA-TLX (Task Load Index)	A widely used tool for assessing perceived workload in human performance studies. The index measures workload across six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration (Hart, 2006)
No User-in-Charge Operator (NUiCO)	Typically, an organisation responsible for responding to vehicle alerts, maintaining, and insuring the vehicle, ensuring safe operation, and managing other activities like toll payments. For example, an organisation that oversees the safe operation of vehicles equipped with a No-User-In-Charge feature - features designed to perform the entire dynamic driving task without a user-in-charge. (Law Commission of England and Wales & Scottish Law Commission, 2022b)

Non-Driving Related Activities (NDRA)	These refer to tasks or activities that a user-in-charge might engage in while the vehicle is in self-driving mode, but which are unrelated to the direct operation or oversight of the vehicle. Activities such as eating popcorn, completing a wordsearch, drinking water etc.
Operational Design Domain (ODD)	The specific area where an automated driving system operates autonomously. This scope can be constrained by factors such as location, time, road type, weather conditions, or other criteria. (Law Commission of England and Wales & Scottish Law Commission, 2022b)
Situational Awareness (SA)	Three levels of responsiveness to the driving environment: Level 1, Perception; Level 2, Comprehension; and Level 3, Projection. These levels correspond to basic awareness, analysis of the current situation, and the ability to predict future outcomes, respectively (Endsley, 1988)
Situational Awareness Global Assessment Technique (SAGAT)	A commonly used measure of SA which involves halting simulated trials and asking participants questions to assess their levels of perception, comprehension, and projection (Endsley, 1998).
Situational Awareness Technique (SART)	A widely used tool for assessing SA across various domains, including aviation, driving, and healthcare. It typically consists of a series of Likert-scale questions designed to gauge an individual's perception and comprehension of their environment, their understanding of the current situation, and their ability to anticipate future events (Salmon et al., 2009).

1. Introduction

1.1 Background to the project

Self-driving technology represents a transformative shift in transport, offering the promise of enhanced convenience and safety, combined with the opportunity to make use of time previously spent driving (Mutzenich et al., 2021). In the early days of automated technology, to standardise the description of automation levels, the Society of Automotive Engineers (SAE) introduced six levels of automation for on-road vehicles (SAE International, 2018). Levels 0 to 2 involves the driver's continuous control, with some assistance from advanced safety systems such as automatic braking. At SAE Levels 3-5, the autonomy increases gradually, with the vehicle capable of handling driving activities for brief periods at Level 3 to being capable of fulfilling all driving at Level 5. The ‘levels’ of automation quickly became shorthand to communicate the shift from human driver to system control but lacked legal clarity.

In 2022, the Law Commission of England and Wales and the Scottish Law Commission (The Law Commissions) published a report suggesting a substantial overhaul of the legal framework governing automated vehicles. This joint report conducted a comprehensive review of the regulatory structure for automated vehicles on public roads and highways and introduced a novel legal entity known as the User-in-Charge (UiC). The Department for Transport (DfT) and the Centre for Connected and Autonomous Vehicles (CCAV) no longer refer to ‘levels’ of self-driving, instead using the Law Commissions’ definition of the UiC: an individual situated within the vehicle and capable of operating the driving controls while a self-driving feature is engaged but not a driver (Law Commissions, 2022a). The UiC is not accountable for any steering, accelerating or braking manoeuvres (referred to as the Dynamic Driving Task (DDT)) and has no obligation to monitor the driving environment or the road when the system operates in self-driving mode. By law, the UiC is no longer the ‘driver’ if the car is in automated mode, but they must possess the necessary qualifications and fitness to assume driving control if the system issues a transition demand. At this point, the UiC becomes the ‘driver’ and assumes full responsibility for taking back control and resuming manual driving (Law Commissions, 2022b).

The first approved self-driving system, under certain conditions, is an Automated Lane Keeping System (ALKS), which pertains to a specific automated system tailored for low-speed scenarios, typically functioning at speeds of 60 kilometres per hour (or 37 mph) or lower (BSI, 2022). ALKS primarily centres on lane-keeping capabilities, managing steering to ensure it stays within its designated lane. ALKS operates independently for prolonged durations without supervision or additional input. According to UN ALKS regulations (2021), when ALKS is activated (currently only allowed on motorways) the UiC must be prepared to take over control if a transition demand is issued by the system.

The transition demand and takeover protocols

A transition demand can be triggered in the case of emergencies, like system failures, or when the circumstances fall outside the specific area (such as location, time of day, road type, weather conditions) where an automated driving system is authorised to operate autonomously, known as its Operational Design Domain (ODD) (Law Commission, 2022). A transition demand could be planned or unplanned: an unplanned event refers to a situation that cannot be predicted beforehand but is considered highly likely to occur, such as changes in speed (slowing down or speeding up), road construction, bad weather, an approaching emergency vehicle, missing lane markings, or debris falling from a truck. In contrast, a planned event is one that is known ahead of time, such as a specific journey point like a highway exit, that necessitates a transition demand (United Nations, 2021).

A transition demand is signalled to the UiC by a **takeover request**, signalled by visual, auditory, or haptic cues, alerting the UiC to situations requiring intervention, such as system limitations or road hazards (Petermeijer et al., 2017). According to UN ALKS regulations (2021), when the ALKS is active, its status must be clearly indicated to the driver via a dedicated optical signal. This signal should feature a clear symbol, such as a steering control or vehicle icon with an "A" or "AUTO," as outlined in UN Regulation No. 121. Additionally, the signal should be placed within the driver's peripheral vision, such as in the instrument cluster or on the steering wheel, ensuring it is easily noticeable.

In the case of a planned event that would prevent the ALKS from operating, the system must issue a takeover request with enough time for the UiC to respond and ensure a safe stop. If the UiC does not respond or there is a critical issue with the ALKS or the vehicle, the system will automatically initiate a Minimum Risk Manoeuvre (MRM) (United Nations, 2021). This MRM will bring the vehicle to a safe stop. If the UiC does not take control, the ALKS will perform the MRM to stop the vehicle before the event occurs (Law Commission, 2022b). According to United Nations ALKS regulations (2021), the UiC has **10 seconds** to take control after the takeover request is issued; if they do not, the system will automatically bring the vehicle to a safe stop. Additionally, any non-driving related activities displayed on the vehicle's screens will be automatically paused as soon as the takeover request is made.

Summary

- The Department for Transport (DfT) and the Centre for Connected and Autonomous Vehicles (CCAV) no longer refer to SAE 'levels' of self-driving, instead using the Law Commissions' (2022) definition of the User-in-Charge (UiC). The UiC does not have to monitor the road or driving environment while self-driving modes are activated but is responsible for taking control when requested.
- The first approved self-driving system, Automated Lane Keeping System (ALKS), operates at low speeds (up to 37 mph) on motorways and requires the UiC to be ready to take over when necessary.
- The UiC must have the qualifications to resume control if a transition demand occurs, shifting from being a passive occupant to an active driver.
- A transition demand occurs when the automated driving system encounters emergencies or conditions outside its Operational Design Domain (ODD).
- Transition demands can be:
    - Unplanned: Unpredictable situations like speed changes, road construction, or bad weather.
    - Planned: Known events like a highway exit that require driver intervention.
- The User-in-Charge (UiC) receives a takeover request via visual, auditory, or haptic cues.

- If the UiC does not respond within 10 seconds, the system initiates a Minimum Risk Manoeuvre (MRM) to safely stop the vehicle.
- Non-driving related activities on the vehicle's screens are automatically paused when a takeover request is issued.

1.2 Literature Review: Situational Awareness in Driving

Definitions

While ALKS is engaged the UiC is not required to monitor the road, yet upon receiving a takeover request, the UiC must suddenly focus on building Situational Awareness (SA) from the surrounding environment to enable a safe takeover. There are various definitions of SA (see (Endsley et al., 2003; Endsley, 2015; Gugerty, 1997, 2011; Lo et al., 2016; Niklasson et al., 2007) but one of the most commonly cited comes from Endsley's original model, which defines SA as,

> "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*" (Endsley, 1988a)

This model divides SA into three levels of responsiveness: Level 1, Perception; Level 2, Comprehension; and Level 3, Projection. These levels correspond to basic awareness, analysis of the current situation, and the ability to predict future outcomes, respectively (Jones & Endsley, 1996). In the context of driving, Level 1 Perception may involve awareness of nearby objects or other road users, Level 2 Comprehension entails understanding and interpreting this information, and Level 3 Projection involves predicting future events such as potential collisions (Mutzenich et al., 2021). Manual driving requires all three levels of SA; perception and comprehension awareness are continuously updated during driving as the environment evolves, and drivers must also anticipate future events, such as what other drivers may do next, or upcoming turns in the road (Endsley, 2020).

The UK Highway Code (2023) emphasises the importance of the Mirrors – Signal – Manoeuvre (MSM) routine as a fundamental practice for safe driving. Before signalling or changing direction, drivers are advised to check their mirrors to assess the position and speed of surrounding traffic. This routine ensures that drivers have a full awareness of their environment, helping to prevent collisions and other hazards on the road. Specifically, Rule 161 of the Highway Code instructs drivers to use mirrors "frequently so that you always know what is behind and to each side of you" (Highway Code, 2023).

Once a takeover request has been issued, the UiC, now acting as the 'driver,' must develop perception of the situation (Level 1 SA) potentially by checking mirrors, comprehend why the takeover request was issued (Level 2 SA), for example by picking up cues in the surrounding environment, and determine necessary future actions (Level 3 SA), such as braking or speeding up, before taking back control.

Time to take over following a takeover request

It stands to reason that drivers who are not looking at the road during periods of automated driving, because they may be legitimately engaged in other non-driving related activities, will need time to develop SA when asked to take back control. Studies have demonstrated that humans possess the ability to visually process a scene within milliseconds, for example, Thorpe et al. (1996) found that participants could quickly grasp the gist of the contents of a scene within 1 second. Although during that time a transitioning UiC/'driver' could grab hold of the steering wheel and apply the brakes, understanding whether that is the correct course of action or could lead to an accident may require more time (Thorpe et al., 1996).

Under the UN regulation for Automated Lane Keeping Systems (ALKS), which specifies that ALKS can be activated under certain conditions on roads where pedestrians and cyclists are prohibited, the UiC has a minimum of **10 seconds** to take control after the takeover request is issued. If the UiC does not take back control, the system will automatically bring the vehicle to a minimal risk condition (e.g. a safe stop) (United Nations, 2021). In the domain of highly automated driving, studies such as Eriksson & Stanton, 2017; Gold et al., 2013, 2016; Lorenz et al., 2014; Melcher et al., 2015; Mok et al., 2017; Radlmayr et al., 2014; Salmon et al., 2012 have all focused on understanding the time required to regain SA after taking back control from automated systems. Yet conclusions are inconsistent regarding how much time is needed to gain necessary SA after periods of self-driving.

Eriksson and Stanton (2017) observed that the average response time for taking back control from a self-driving system during on-road driving was approximately 3 seconds. According to Mok et al. (2017), it typically takes between 5 and 8 seconds to regain control of an autonomous vehicle after being engaged in an active secondary activity, such as playing on a tablet.

In contrast, Lu et al. (2017) conducted an experiment where participants watched videos of varying lengths and were then asked to reproduce the traffic layout of a three-lane road they had seen. Their findings suggested that it takes between 7 and 20 seconds to develop sufficient Level 1 SA (Perception) to successfully complete this perceptual activity (Lu et al., 2017). However, when participants were tasked with assessing the relative speeds of other vehicles in relation to their own, it took them 20 seconds or more. Achieving full awareness, especially for complex activities involving comprehension (Level 2 SA) and projection of future events (Level 3 SA), may take more time than simple perceptual awareness (Level 1 SA).

In simulated scenarios, such as automated driving in a driving simulator, it typically takes participants approximately 12 seconds to feel sufficiently secure to resume manual control of driving (Coster, 2015). Vogelpohl et al. (2018) suggests that while drivers can deactivate automation relatively quickly, they may require more time to build SA and respond effectively to unexpected traffic events (Vogelpohl et al., 2018). Developing good SA is critical for anticipating the actions of other road users in complex traffic scenarios, especially after drivers have deactivated automation (Wulf et al., 2015).

These studies highlight that quickly taking control does not guarantee SA of the road environment. Therefore, a comprehensive measure of SA post takeover should include all three levels to provide a valid assessment of readiness and capability of the UiC to safely manage the vehicle after a transition.

Summary

- Once a takeover request has been issued, the UiC, now acting as the 'driver,' must take back control within 10 seconds.
- How long it takes the UiC to gain SA after a takeover request is critical to ensuring a safe transition.
- Humans can quickly process the gist of a scene, but complex activities like resuming control of a moving car may require more time to develop SA (Thorpe et al., 1996).
- Response times to take over control with sufficient SA vary, range from approximately 3 to 20 seconds in different scenarios (Eriksson and Stanton, 2017; Coster, 2015; Lu et al., 2017).

Non-Driving Related Activities (NDRAs) and Situational Awareness (SA)

Periods of non-driving, which will occur whilst using ALKS technology, will allow drivers to disengage and re-engage with the driving activity (Radlmayr et al., 2014). As self-driving technologies become more prevalent, questions arise regarding the permissibility of engaging in Non-Driving Related Activities (NDRA) while using vehicles with activated self-driving features such as ALKS. There is potential for certain non-driving activities to be allowed, provided they do not compromise the driver's ability to resume control of the vehicle safely when a takeover request is issued by the system.

It is not yet clear what activities people might choose to engage in while the ALKS is activated, but observations from previous research into simulated self-driving offer some insights. In studies where participants were allowed to bring their own items to engage with during periods of automation, most participants brought and used mobile phones, tablet computers, laptops, books, and printed papers to occupy themselves during automated driving (Large, Burnett, & Salanitri, 2019; Shaw et al., 2020). Additionally, it's likely that people may choose to eat or drink while the system is active, even though these activities are currently prohibited by the Highway Code, they are commonly practiced (Department of Transport, 2023).

The UK Government’s consultation on ALKS regulations, with responses from local transport authorities, technology developers, trade associations, legal firms, academia, insurers, and manufacturers, revealed mixed responses on which NDRAs should be allowed while the self-driving features are activated (CCAV, 2021). There was no clear consensus on phone use, with opinions divided between allowing all phone activities and restricting use to hands-free calls. Many respondents expressed concern that engagement in any activities would be too immersive and questioned whether 10 seconds would be sufficient time to respond to a transition demand if performing other activities. There was a strong sentiment amongst respondents that the UiC should remain alert, with eyes on the road, ready to take back control at any moment (CCAV, 2021). However, the Law Commission report determined that the UiC will not be expected to keep their eyes on the road when the system is in ALKS or another automated driving feature, otherwise there would be little attraction in self-driving features (Department for Transport, 2022).

Effect of NDRAs on Situational Awareness following a takeover request

ALKS technology is currently only permitted on UK public roads with strict limitations, so driving simulator studies are used to understand their impact as fully self-driving systems. These studies allow researchers to explore scenarios and driver interactions that cannot yet be tested on public roads.

Research on how activities performed while driving automation is engaged affect SA following a takeover request has produced mixed results. Some studies have found that NDRAs have had a negative effect on SA, for example, Vogelpohl et al. (2018) examined the impact of distractions on the formation of SA, measured by the time to first gaze to the side mirrors and speedometer after taking over from automation. Participants were assigned activities that included playing games requiring high cognitive engagement, which involved fast-paced visual and motor responses, and reading text or articles on a screen or tablet. The study found that participants in distracted automation conditions exhibited a delayed formation of SA compared to those in no-activity conditions or manual driving groups. Marti et al. (2022) showed that being deeply involved in a task just before a takeover request significantly reduces SA and impairs safe vehicle control when starting manual driving. Participants experienced varying levels of distraction until the moment of the takeover request, ranging from no distraction to being fully engaged in a non-driving task (reading aloud a scrolling text on a side tablet) (Marti et al., 2022). Drivers who were completely immersed in the non-driving activity had more collisions after an unplanned takeover compared to those with no distractions. This evidence suggests that distractions during automation driving significantly impair the ability to form SA when asked to takeover.

However, some studies have found that engaging in a non-driving activity during periods of automation can have a positive effect on SA as participants are kept in a state of arousal even when not expected to be monitoring the road. Miller et al. (2015) found that participants who watched videos or read on a tablet were less likely to show drowsiness (6%) compared to those supervising automated driving (27%). Du et al (2020) gave drivers in a driving simulator a hands-free activity of watching a film on a tablet on the dashboard, before a takeover request was issued. They found that without the need to physically end the activity and put down the device, drivers could immediately switch their attention from the tablet to the driving scene, taking less than 2 seconds to transfer attention. Jiang, Wang, & Tang (2024) found that gamified attention activities during simulated automated driving scenarios can improve reaction times to takeover and decision accuracy. They tested two types of games: ‘Situational’ games like ‘Fruit Ninja’ and ‘Subway Parkours’, which are fast-paced and require undivided attention, and ‘Non-situational’ games like ‘Tetris’ and ‘2048’, which are slower and less demanding. SA was measured by assessing takeover time, acceleration, and time to collision. Both gaming groups demonstrated faster takeover times compared to the monitoring group, suggesting higher arousal levels. In conclusion, it could be that rather than being distracting, integrating gamification into attention activities could enhance SA and takeover quality, potentially improving safety in automated driving (Jiang et al., 2024).

Types of Non-Driving Related Activities and Situational Awareness

Understanding the effect of different types of tasks is crucial, as the specific NDRA the driver chooses to occupy their non-driving time could have an impact on how much time it takes them to build up sufficient SA to begin driving again safely. Radlmayr et al. (2014) used the

Surrogate Reference Task (SuRT) and n-back activity to mimic the visual and attentional demands of real-world activities on driving. The Surrogate Reference Task (SuRT) is a standardised visual task used to simulate cognitive and visual workload. Participants are asked to identify and select a target (usually a small circle) among distractors on a screen while performing another task, such as driving. The n-back task is a cognitive exercise where participants must identify when the current stimulus matches one presented "n" steps earlier in a sequence, challenging their working memory and attention. They found no significant effect of either NDRA on the takeover process. However, this activity may not fully represent the variety of real-life NDRAs encountered in automated vehicles as it is a meaningless yet difficult activity, rather than an entertaining one.

More realistic NDRAs can be broadly categorised into those that involve technology, such as using mobile phones, which can differ greatly depending on whether the device is handheld (requiring button presses and key inputs) or mounted (such as watching a film). Other tasks may not involve technology but still engage cognitive and visual modalities, such as completing a crossword puzzle, doing a word search, or reading. Additionally, some NDRAs are more motor based, like eating or drinking. Different categories of NDRA may have distinct implications for how quickly and effectively a UiC can respond to a takeover request.

For instance, Chen et al. (2023) studied the effects of work-related and entertainment activities on mobile phones during automated driving and found that entertainment activities led to quicker takeover times than work tasks, indicating that tasks requiring higher cognitive engagement take longer to recover from. Gold et al. (2015) similarly found that drivers engaged in cognitive activities performed better during unplanned takeovers compared to those performing visual-motor tasks.

On the other hand, Dogan et al. (2019) found no significant difference in takeover performance when drivers engaged in activities like writing emails or watching videos during planned takeovers at low speeds. Vogelpohl et al. (2018) observed that while automation itself affected preparation for action, the type of NDRA did not significantly impact takeover time or brake reaction time when using representative tasks like reading or playing games. Zeeb et al. (2016) also found that while secondary activities did not significantly affect the time to return hands to the steering wheel, they did deteriorate takeover quality, especially for tasks like reading or watching videos, where participants deviated more from the lane centre.

Finally, Jenness et al. (2002) assessed various activities such as eating, operating a CD player, and using voice-activated dialling. The eating task was operationalised as unwrapping and eating a cheeseburger. They found that reading and operating the CD player led to the highest number of lane-keeping errors and glances away from the road. In contrast, eating and voice-activated dialling, although still distracting, resulted in fewer errors.

These research findings highlight that different NDRAs demand varying levels of cognitive, visual, and motor engagement, which can significantly influence a driver’s ability to promptly and effectively resume control of the vehicle following a takeover request. However, the research is mixed, as some studies suggest that certain NDRAs can actually result in better takeover performance, indicating that the type of task and the context in which it is performed are critical factors in determining its impact.

Summary

- Research on the impact of NDRAs on SA and takeover performance shows mixed results, with some studies finding negative effects (e.g., Vogelpohl et al., 2018; Marti et al., 2022) while others suggest potential benefits, such as maintaining arousal and improving reaction times (e.g., Miller et al., 2015; Du et al., 2020; Jiang, Wang, & Tang, 2024).
- NDRAs can be categorised into technology-based activities, like using mobile phones or watching films, and non-technology tasks, like reading or motoric tasks such as eating, each with different implications for how quickly and effectively a driver can respond to a takeover request (Chen et al., 2023; Jenness et al., 2002).
- Various studies have found that cognitive tasks can enhance takeover performance (e.g., Gold et al., 2015), while visually demanding activities tend to impair takeover quality, particularly in maintaining lane position and response accuracy (e.g., Zeeb et al., 2016).
- The overall effect of NDRAs on takeover performance is highly dependent on the specific type of task and the context in which it is performed, underscoring the need to carefully consider these factors when evaluating their impact on driving safety (Dogan et al., 2019; Vogelpohl et al., 2018).

1.3 Metrics to measure Situational Awareness after a takeover request

One commonly used measure of SA is the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988b), which involves pausing simulations to ask participants questions assessing their perception, comprehension, and projection (Endsley & Jones, 1996). While SAGAT has been widely used in fields like military aviation, it has been criticised for not fully capturing the dynamic perceptions of drivers in real-time environments. This method may also limit participants' SA by focusing narrowly on specific aspects of awareness, making it less applicable in complex driving scenarios where a UiC must respond to a takeover request (Mutzenich et al., 2021). There has been no consistent approach to measuring SA during takeover scenarios, with different studies exploring alternative methods for assessing SA during takeovers in automated driving. The following section outlines those methods, including performance measures, behavioural observations, eye-tracking techniques, physiological indicators, and subjective assessments.

Performance measures

Performance measures of driving consist of parameters such as position on the road (both lateral and headway distance to cars around them), speed, acceleration (both lateral and longitudinal), as well as steering wheel angle and rate (Gold et al., 2015; Louw et al., 2015; Merat et al., 2014). These metrics are derived by computing the minimum, maximum, mean, or standard deviation of the gathered data points (Shariati et al., 2023).

Studying the standard deviation of steering wheel angle and/or lateral lane position post-take over allows researchers to infer participants' SA of the driving environment following a takeover request (Shariati et al., 2023). In typical manual driving, drivers make small, continuous adjustments to their steering wheel to maintain their lane position as conditions on the road change. However, when the workload increases, such as during complex manoeuvres or high traffic situations, drivers tend to make fewer steering corrections, leading to larger deviations from their intended lane position (Eriksson & Stanton, 2017). Erikson (2017) found that the standard deviation of steering input increased significantly when participants resumed control from automated driving compared to a manual baseline. Using this metric can provide insight into the level of cognitive and physical demand placed on the driver during the takeover in response to different driving scenarios.

Poor initial takeover performance immediately after resuming driving control can also be used to infer SA is undeveloped after takeover. Burnett et al. (2019) demonstrated in a driving simulator study, that drivers who spent a week using an automated vehicle for their daily commute engaging in a range of immersive, non-driving related tasks during periods of self-driving, were unprepared when required to resume active control. Participants showed high levels of swerving and speed variability during the 10 seconds of manual driving immediately following a scheduled handover. These signs of instability and variability suggest suboptimal SA through a lack of readiness to handle the driving environment post takeover (Burnett, Large, & Salanitri, 2019).

Many studies have used the metric of time to as a key measure to infer SA during the transition from automated to manual driving (see Gold et al., 2013, 2016; Lorenz et al., 2014; Melcher et al., 2015; Radlmayr et al., 2014; Salmon et al., 2012; Mok et al., 2017; Eriksson and Stanton, 2017; Lu et al., 2017; Coster, 2015; Vogelpohl et al., 2018, Dogan et al., 2017; Du et al., 2020). For example, Hands-on Reaction Time (HRT), measures the time from the initiation of the takeover request to when the driver places their hands on the steering wheel. Vogelpohl et al. (2018) measured the time to takeover by how quickly the participant responded to a takeover request by grabbing the steering wheel. SA was measured by the time it took for the driver to gaze back to the centre of the road. Engagement in an NDRA significantly impacted both take-over time and gaze patterns, leading to poor quality of take over (Vogelpohl et al., 2018). While a quick HRT indicates that the driver has rapidly taken control of the vehicle, it does not necessarily mean that they have achieved sufficient SA. Quickly placing hands on the wheel without adequately assessing the driving environment could lead to poor decision-making and unsafe driving, highlighting that HRT alone is not a comprehensive measure of takeover effectiveness.

Time to collision is also used as a measure of how urgently the driver needs to react, with lower values suggesting a delayed response in initiating lane changes or braking actions, which can also indicate a lack of SA post-takeover. Examining whether drivers have collisions with objects or other vehicles can also tell us about the amount of SA that the UiC has on taking back control, as awareness of objects around you is part of Level 1 SA Perception (Radlmayr et al., 2014).

Gold et al (2015) looked at HRT, time to takeover and takeover quality as measures of SA following a takeover request. They defined reaction time as the time between the take-over request and the first gaze directed away from the non-driving task and time to take over as when the participant first starts carrying out a manoeuvre (turned the steering wheel and/or braked). Take over quality was measured by assessing how much participants sped up after taking control and how long it took them to hit another object or car (time to collision). The lower the acceleration and the longer the time to collision can represent a safer handling of the situation and therefore better quality take over. They found that drivers reacted faster in manual mode than under the presence of the automated system.

Zeeb et al. (2016) measured the time it takes for drivers to regain control after a take-over request and when they shift their gaze back to the road. Eye-tracking technology and vehicle data was used to observe when drivers shifted their gaze from NDRAs to the road and took control of the vehicle. SA was measured by examining how quickly drivers could perceive the driving environment and either accelerate, steer or brake once their focus returned to the road. When drivers were distracted, they took longer to respond, which led to poorer performance in taking control of the vehicle (Zeeb et al., 2016).

Although when ALKS is engaged the UiC does not need to monitor the driving environment, ‘drivers’ may continue to make checks to the road while carrying out other activities, as they may not have full automation trust. Feldhütter et al. (2017) measured SA after a takeover request by focusing on how long and how often participants looked at the driving scene during automated driving. Longer automation periods led to slower take over time, as participants glanced at the driving scene less, affecting participant’s take over quality (Feldhütter et al., 2017). Their research suggests that as automation periods become longer, checks on the road reduce but so does takeover time as the UiC is fully immersed in the NDRA.

These studies primarily focus on measuring the time to take over but indirectly assess SA by evaluating how quickly drivers perceive their environment and update their understanding after refocusing on the road. The results indicate that engagement in NDRAs significantly affects time to take over and gaze patterns, leading to delays in responses. While time to take over measures driver responsiveness, it does not guarantee awareness.

Summary

- Standard deviation of steering wheel angle and lateral lane position post-takeover can provide insights into User-in -Charge SA post-takeover by reflecting their ability to maintain control during complex driving scenarios.
- Many studies have used the metric of time (Hands-on Reaction time, time to collision) to measure driver responsiveness, but quick takeovers do not guarantee effective SA.
- Takeover quality shown by initial takeover performance, such as swerving and speed variability, can indicate low SA, particularly when drivers struggle to adjust immediately after resuming control from an automated system.

Behavioural measures

Behavioural observation techniques, often using in-cabin cameras and thematic video analysis, have been instrumental in understanding the takeover process. These observational methods allow researchers to observe critical aspects such as hand positioning on the steering wheel, the extent of interaction with non-driving activities, and the overall effectiveness of the driver’s disengagement from non-driving activities and transition back to manual control.

Large et al., (2019) investigated the behaviours that drivers engaged in when self-driving features were active for long periods of time. Over 5 consecutive days, experienced drivers performed a simulated 30-minute motorway journey resembling their daily commute in a self-driving vehicle. Participants were encouraged to bring personal objects or devices and freely interact with them during the automated period of driving. Thematic video analysis revealed that participants quickly engaged in various secondary activities that demanded significant visual, manual, and cognitive attention, including adjustments to their seating posture. The steering wheel often served as a support for these activities meaning that when an unplanned takeover request occurred, drivers had to abruptly abandon these tasks, indicating the challenges of quickly shifting focus back to the driving task (Large, Burnett, & Salanitri, 2019).

A safe takeover heavily depends on the ability of the UiC to fully disengage from other activities and re-engage with the driving task upon receiving a takeover request. Burnett et al. (2019) highlighted concerns that drivers often used the 60-second "prepare-to-drive" notification given as a takeover request in their driving simulator study not as a cue to actively re-engage with driving but rather as a signal to start wrapping up their secondary activities. This behaviour delayed their readiness to assume full control of the vehicle and failed to provide sufficient SA needed for safe manual driving. Shaw et al. (2020) examined driver behaviour across different driving phases, including the transition from automated to manual control. They found that many drivers, after receiving a takeover request, briefly glanced at the road but quickly reverted to the activity they were engaged in prior to the takeover request instead of fully preparing to drive. Incomplete disengagement from secondary tasks and insufficient re-engagement with the driving task can compromise safety, as drivers attempt to split their attention between the non-driving activity and the driving task.

Summary

- Behavioural studies using in-cabin cameras and thematic video analysis help researchers observe key aspects of the takeover process, such as hand positioning, interaction with non-driving tasks, and the effectiveness of disengagement and re-engagement.
- Disengagement from non-driving activities is crucial for a safe takeover. UiC must fully shift their attention from secondary tasks to the driving task upon receiving a takeover request.
- Studies have found that participants revert back to their non-driving activities after takeover request rather than preparing for driving, leading to delayed readiness and insufficient SA (Burnett et al., 2019, Large et al., 2019).

Eye tracking measures

Eye tracking has emerged as a critical tool for understanding how drivers allocate visual attention and build SA during and after a takeover request in automated driving scenarios. By examining where and how long drivers focus their gaze, researchers can infer the quality of SA and the driver's readiness to safely resume control of the vehicle.

Research by Kunze et al. (2019) demonstrated that drivers who exhibited shorter fixations in a peripheral search task during the 40 seconds before a takeover request tended to have higher SA scores. This suggests that the ability to quickly scan the environment upon receiving a request to takeover without lingering on specific areas may contribute to better SA. Liang et al. (2021) found that greater gaze dispersion and more time spent looking at the road scene on takeover were positively correlated with SA scores. Evidence also shows that previous engagement in visually demanding NDRAs, such as watching films, impairs SA after the takeover request, highlighting the importance of distributing visual attention effectively to maintain a high level of SA (Du et al., 2020, Liang et al. 2021).

Eye tracking studies have also focused on where the driver looks immediately after a takeover request. For instance, Vogelpohl et al. (2018) analysed drivers' reactions following a takeover request during an automated drive with highly distracting NDRAs. They found that distracted drivers showed a delay of up to 5 seconds in their first glance at the side mirror or speedometer, compared to drivers in no-activity conditions. Although these drivers were relatively quick in physically taking back control, their delayed visual attention to critical driving cues suggested a slower development of SA, potentially compromising safety.

Eye tracking metrics can help evaluate how NDRAs impact drivers' monitoring behaviour and SA during different phases of driving, including automated and manual control. Eye tracking measures appear to indicate a longer time is needed to regain SA after transitioning from automation to manual control than suggested by the studies discussed earlier in Section 1.3 (e.g., Eriksson and Stanton (2017), Mok et al. (2017), Coster, 2015, Lu et al, 2017). For example, Merat et al. (2014) conducted a study to explore drivers' ability to resume control from highly automated vehicles under two conditions: one where automation disengaged at regular intervals and another where transition to manual control was triggered by the duration of drivers looking away from the road. Using eye tracking data, the study examined visual attention patterns as drivers resumed manual control in both scenarios revealing that drivers exhibit more dispersed horizontal gaze patterns during autonomous driving compared to manual operation. The findings revealed that drivers required approximately 40 seconds to regain sufficient and stable control of driving from automation. This suggests that even if a takeover occurs, drivers may still be building up SA and may not yet be ready to drive safely, even if they have the operational capacity to take control.

The frequency and duration of mirror checks are important indicators of how well drivers are re-engaging with the driving task after a takeover request. Monitoring where drivers direct their gaze—towards mirrors, the road ahead, or back to the NDRA—provides insights into their level of engagement and SA. UN ALKS regulations (2021) stipulate that driver attentiveness can be inferred from their gaze direction, whether towards the road, mirrors, or other driving-related cues. Yu et al. (2023) found that drivers directed a significant percentage of their gaze towards the central instrument panel within the first 6 seconds after a takeover request, which could indicate efforts to quickly regain SA by searching for information.

Pupil diameter is another eye tracking metric used to assess cognitive load and SA. In automated driving scenarios, changes in pupil diameter can be a reliable indicator of the driver's cognitive state and readiness to take over, especially after being engaged in an NDRA. Ahlstrom & Friedman-Berg (2006) noted that larger pupil diameters indicate higher cognitive load, as the brain processes information more intensely. An increased pupil diameter change rate typically signals heightened cognitive load or stress, which can occur when a driver is trying to regain full awareness of the driving environment after a period of automated driving. Conversely, a lower change rate might indicate that the driver is more relaxed or less engaged.

How visual attention is allocated during and after a takeover request, and the subsequent visual behaviour and attention distribution are crucial for rebuilding the SA necessary for safe driving. Being immersed in NDRAs can significantly impair this process, delaying the time it takes to regain full SA after a takeover request.

Summary

- Eye tracking studies have shown that visual attention allocation, such as shorter fixation times and systematic scanning of the environment immediately after a takeover request are associated with higher SA and better takeover performance.
- Metrics like pupil diameter and mirror checks provide additional insights into cognitive load and engagement with the driving task, further informing SA levels during and after a takeover.
- Engaging in visually demanding NDRAs before a takeover can delay the formation of SA, potentially compromising the safety and effectiveness of the takeover.
- Studies indicate that drivers who are visually distracted by NDRAs may physically take control quickly but require more time to rebuild the necessary SA for safe driving.

Physiological measures

Electroencephalography (EEG) is a lesser used but promising tool for assessing SA during and after a takeover request in automated driving scenarios. EEG measures brain activity, providing insights into cognitive processes and workload levels, which are crucial for understanding how drivers regain SA after a takeover.

In a study by Van Miltenburg et al. (2022), EEG was used to monitor brain activity in drivers engaged in NDRAs of varying difficulty levels, before and after the moment they were required to take control of an automated vehicle. The study found that EEG data could reflect participants' levels of distraction, although it was not always a consistent predictor of driving performance during the takeover. This suggests that while EEG can provide useful information about cognitive load, it may not always directly correlate with the effectiveness of the takeover.

There is, however, evidence that specific brain signatures associated with SA can be detected using EEG. Novel research by Kastle et al. (2021) identified distinct activation patterns in the occipital and temporal lobes—areas of the brain linked to visuo-spatial ability, memory, and reasoning. This study also attempted to correlate EEG signals with SA by categorising SA based on reaction time functions, with EEG signals labelled as indicating either ‘poor’ or ‘good’ SA depending on whether the reaction time was above or below a certain threshold (Kastle et al., 2021). This method, developed using an EEG dataset provided by Cao et al. (2019), suggests that EEG can be used to differentiate between varying levels of SA in real-time during a takeover event. The study highlighted that EEG traces could potentially reveal whether a driver has successfully regained SA during a takeover, making it a valuable tool for assessing cognitive readiness (Cao et al., 2019).

These findings indicate that EEG can be an objective measure of SA, particularly when combined with other data sources like eye tracking to provide deeper insights into how drivers regain SA and ensure safe driving after a takeover request.

Summary

- Studies have shown that EEG measurements of brain activity can reflect participants' levels of distraction during NDRAs but may not always predict driving performance during a takeover (Van Miltenburg et al.,2020).
- EEG, when combined with eye tracking data, can objectively measure SA, via specific brain activation patterns in the occipital and temporal lobes associated with visuo-spatial ability and memory (Kastle et al. (2021).
- Researchers have correlated EEG signals with reaction time to categorise SA levels, indicating that EEG can differentiate between ‘poor’ and ‘good’ SA during a takeover.

Subjective measures

NASA-TLX (Task Load Index) and SART (Situation Awareness Rating Technique) are commonly used subjective measures in research to assess cognitive workload and SA, particularly in automated driving scenarios.

The NASA-TLX is a widely recognised tool used to assess perceived workload across six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration (Hart, 2006). Participants rate these dimensions on a Likert scale, helping researchers quantify the cognitive and physical demands experienced during tasks, such as driving or engaging in non-driving activities during automated vehicle operation.

Heo et al. (2022) used NASA-TLX to evaluate mental workload under various environmental conditions during simulated driving and found that different weather conditions significantly impacted workload, particularly when drivers had to respond to takeover requests (Heo et al., 2022). Jiang et al. (2024) employed NASA-TLX to assess the workload of participants engaged in situational and non-situational games during automated driving. The results indicated that situational games improved SA and takeover performance while maintaining lower workload levels compared to non-situational games. Eriksson & Stanton (2016) measured workload using NASA-TLX after drivers engaged in secondary tasks during automated driving. Workload scores varied depending on the task and the timing of the takeover request (Eriksson & Stanton, 2016).

SART is a subjective tool designed to assess SA by evaluating how participants perceive their ability to allocate attention, understand the current situation, and predict future events. It uses a series of Likert-scale questions to gauge an individual's SA in various contexts, including driving and aviation. SART has been used to evaluate the impact of different driving scenarios on SA, revealing that complex driving tasks and challenging driving conditions can significantly reduce a driver’s perceived SA (Li et al., 2023).

Summary

- Both NASA-TLX and SART provide valuable insights into how drivers perceive their cognitive workload and situational awareness during automated driving.
- NASA-TLX is effective in quantifying workload during various tasks, while SART offers a direct measure of SA, helping researchers understand how well drivers can maintain or regain awareness after a takeover request.
- Studies utilising these tools highlight the variability in SA and workload depending on the type of non-driving activity, environmental conditions, and task complexity, emphasising the need for careful consideration of these factors in the design of automated driving systems.

1.4 Brief and Project Aims

This project focused on exploring the implications of NDRAs in vehicles with self-driving capabilities, when a UiC must respond to a transition demand. As the use of ALKS becomes more prevalent, it is essential to understand which NDRAs can be performed (within the 10 seconds mandated by ALKS regulations) without compromising the ability to safely resume control of the vehicle.

The project had the following objectives:

- Investigate which non-driving related activities (NDRAs), if any, can be safely performed during non-driving periods in cars with self-driving features but requiring a UiC.
- Establish mechanisms for measuring SA and determine appropriate thresholds to ensure safe takeover and manual driving resumption.
- Understand potential variations in the impact of NDRAs across different scenarios to inform policy development.

To achieve these objectives, Lacuna Agency worked in partnership with University College London (UCL) and Loughborough University (LU), combining our expertise in Human Factors research and driving simulation. The project used a simulator-based approach to closely replicate real-life driving conditions. This collaboration, together with DfT and CCAV aimed to provide evidence-based insights that will inform future policy development and contribute to safer implementation of automated driving technologies.

2. Method

2.1 Participants

A representative sample reflecting the demographics and characteristics of the general UK population was collected for this project, with specific restrictions on participants related to glasses wearing and driving license requirements. The objective was to ensure that the sample accurately represented the broader UK population in terms of factors such as age, gender, socioeconomic status, ethnicity, and other relevant characteristics. A total of 97 participants were recruited, split between two UK locations: an urban setting in the South-East and a semi-rural location in the Midlands, to capture a cross-section of UK drivers. Testing was conducted at both University College London (UCL) and Loughborough University (LU) in Leicestershire to analyse data from both urban and semi-rural areas. This approach was intended to ensure that the data reflected diverse driving conditions and environments, as self-driving vehicles are a public-facing issue that must represent the public effectively.

Out of the recruited participants, 66 completed the study at UCL and 31 at LU. The sample included a wide age range from 18 to over 66 years, with 47 males and 40 females, and a diverse mix of ethnicities and socioeconomic backgrounds. Participants were compensated £90 for their travel and time, with the study taking approximately one and a half hours to complete. Recruitment methods included both internal and external databases, local networks, and social media groups. Additionally, a snowball sampling technique was employed, where participants were encouraged to refer others who met the study criteria. After accounting for non-compliance, simulator sickness, or data loss, the final dataset analysed included 87 participants (63 from UCL and 24 from LU).

Sample restrictions

The sample for this study had several key exclusion criteria to ensure both participant suitability to the project aims and relevance to the type of individuals likely to drive a car with ALKS technology. All participants needed to have normal vision, as they were required to drive, read, watch a film, and use a mobile phone without glasses since they would be wearing eye-tracking glasses. Additionally, all participants held a full UK driving license and resided in the UK. The sample represented a diverse range of driving behaviours, including variations in transport use, driving experience, and frequency of driving. A clean driving record was mandatory, meaning participants could not have any speeding fines. To avoid bias of involvement through support of new technology, up to 20% of the sample were classified as technology rejectors, ensuring that individuals who were sceptical of new driving technologies were also represented in the study.

Individuals who had participated in similar research, such as driving simulators, within the past six months were excluded to prevent participation purely for incentives. Those prone to car sickness or who felt nauseous while reading or watching films in a car were also excluded as they were likely to be unable to complete the study due to simulator sickness. Medical conditions that could impact participation (including epilepsy, seizures, migraines, adverse reactions to bright or flashing lights, motion sickness, sensitivity to light, and other neurological disorders) were excluded. Additionally, participants taking medication that might impair cognitive function or reaction time were also not eligible.

2.2 Testing environment

UCL testing took place in a high-tech driving simulator at the Intelligent Mobility Group (IM@UCL) located at the PEARL lab in Dagenham East, London, while testing at LU was conducted in the Human Factors Research Lab at Loughborough University using a mid-fidelity driving simulator.

Driving simulators

At UCL, the simulator used a black modified subcompact crossover SUV for the study, connected to an advanced computer system with the engine and fluids removed. The interior and exterior mirrors have been replaced with screens that display the simulated driving environment, making them fully functional. The dashboard operates like a real car, showing speed and engine data. An additional screen on the centre console enables communication between the vehicle and the driver. In front of the vehicle is a 180-degree curved screen with a 2.25-meter radius is used to project the simulation. This is achieved with three projectors, creating a high-resolution display of 12288×2160 pixels at 30 Hertz, seamlessly merged and warped to match the screen’s curvature. The vision system offers a 180° horizontal field of view, while the motion and driver system is provided by Ansible Motion Ltd., featuring a Hand Wheel Loading System and pedal sensors. The entire simulation is Unity-based, with integrated audio for a comprehensive driving experience.

View of the UCL simulator 's exterior showing the rear of the vehicle and the simulated scene of the road ahead on a 180-degree curved screen.

At LU, the driving simulator features a Ford KA car placed within a specialised environment composed of three large screens that form an open cube around the front, left, and right sides of the vehicle. Each screen measures 3200mm in width and 2451mm in height, displaying high-resolution visuals at 1920x1200 pixels and 120Hz, using ultra-short throw projectors. The rear-view and side mirrors are digitally projected onto the screens in their respective positions, providing an accurate representation of the vehicle's surroundings. Inside the vehicle, a digital speedometer, positioned in the standard location, displays the vehicle's speed.

View of the LU simulator 's exterior showing the rear of the car and the side of the car and the three sides of the simulated scene of road ahead. The projection of the rear-view mirror image on the screen can be seen ahead of the vehicle.

Both simulators are designed to provide real-world driving sounds from the car and surrounding traffic. The driver interacts with the vehicle through traditional controls, including the throttle pedal for acceleration, the brake pedal for deceleration, and the steering wheel for initiating turns, offering a realistic and responsive driving experience.

Eye tracking

Both testing sites used Tobii Pro eye-tracking glasses; UCL used Tobii Pro 3, while LU used Tobii Pro 2 to measure eye movements. Eye movements were recorded with a 50Hz frame rate, to ensure accurate tracking of fast eye movements.

EEG (UCL only)

A 32-channel EEG system (g.tec model g.NautilusPro) with gold-plated dry electrodes (type g.Sahara) was used. The electrodes were placed on participants' scalps according to the international 10-20 system, a standard for electrode placement in EEG studies. The data was recorded at a sampling rate of 250 Hz, with a high-pass filter set at 2 Hz to avoid drift and a notch filter at 48−52 Hz to exclude interference from power lines. This setup was chosen to ensure precise measurement of brain activity during critical moments, such as when participants received a takeover request and began to regain control of the vehicle. Four out of 63 participants at UCL were unable to wear the EEG cap due to their hairstyles, which prevented proper contact between the electrodes and the scalp. Hairstyles that include thick braids, large buns, or other styles that create significant gaps between the scalp and the EEG cap can interfere with the electrodes' ability to detect electrical signals from the brain.

While UCL uses a 'dry' EEG system, LU was equipped with a 'wet' EEG system, which requires the application of electrode paste to the participant’s scalp to improve conductivity. The 'wet' EEG was deemed too invasive for this study, as it involves a more complex setup process and can cause discomfort for participants, potentially affecting their performance and the study's outcomes.

In addition to technical considerations, there were procedural differences between UCL and LU that influenced the use of EEG. At LU, the transition from automated to manual control required manual input from researchers and the participant's verbal confirmation, "ready to drive." In contrast, UCL's simulator automatically disengaged automation when participants assumed control through steering or pedal input. Verbalising "ready to drive" at UCL could interfere with EEG signal quality by introducing speech-related artifacts, which could obscure the brain activity being measured. Therefore, only UCL collected EEG data in this study. The EEG findings are recorded in a separate report.

A participant in the UCL simulator, reading a magazine while completing one of the NDRA activities. The participant is wearing an EEG cap and eye-tracking glasses.

GoPro recordings

Findings from studies above suggest that when a takeover request is issued, participants' responses and interaction with the NDRA can vary. Some may immediately disengage from the activity and focus on the driving environment, while others may take additional time to complete their current activity before turning their attention to the road. Video recordings of all trials for all participants were made by using a GoPro HERO 7 camera (1080p, 30Hz) were used positioned to the rear and left of the driver seat on the ceiling. This gives a view of the driver from behind, which also captured the centre console and NDRA.

The use of in-cab cameras allowed the research team to observe these types of behaviours in real-time. It provided insights into how participants prioritise between the activity at hand and the need to assume manual control of the vehicle and whether some NDRA are more demanding of attention than others.

2.3 Experimental Design

The focus of the study was to assess how participants responded to a takeover request from a simulated self-driving vehicle after a period of engaging in various non-driving related activities. The study used a within-participants design, meaning that each participant completed all the trials, and their performance was compared across different trials rather than with other participants. This approach was chosen to minimise individual differences between participants, for example in driving ability.

The study featured two motorway-based scenarios:

> **Scenario 1 “Roadworks”:** Participants had to take over manual control and slow down after encountering roadworks signs.
>
> **Scenario 2 “Congestion”:** Participants had to take over control as dense traffic cleared and then speed up.

In each scenario, participants completed a baseline activity where there was no non-driving task, but they still received a takeover request. They also completed three categories of NDRAs randomised across all scenarios: motor activities, cognitive and visual activities involving technology, and cognitive and visual activities without technology.

Scenario Design

The scenario design for this study aligns with ALKS regulations, which require that the system operates within its specified Operational Design Domain (ODD), in this case, fair weather conditions. Two primary scenarios were developed to examine the impact of NDRAs on participants' ability to respond to a planned (non-critical) takeover request: roadworks and congestion. Both scenarios are designed to trigger a planned takeover request due to changes in speed, simulating situations where the ALKS system encounters conditions beyond its operational limits. These scenarios were selected to assess participants' ability to manage the transition from automated to manual control in realistic yet challenging driving situations, providing insights into the effectiveness of current automated driving systems and the critical role of human intervention in ensuring safety.

Roadworks scenario: The ‘ego vehicle’ (the car the participant is in) drives autonomously at 68 mph in light traffic, staying in the left-hand lane. After 2-4 minutes, a Roadworks sign appears, prompting a takeover request approximately 100 metres (3-4 seconds) before the sign. Additional Roadworks warnings follow for context. About 20 seconds after the takeover request, an overhead gantry displays a mandatory 50 mph speed limit, causing other traffic to slow down. A roadside sign then indicates Roadworks 0.5 miles ahead. Upon taking manual control, the participant is expected to decelerate in response to speed signs.

Congestion scenario: The ego vehicle drives autonomously at 37 mph due to traffic congestion, surrounded by dense but not fully stopped traffic. The vehicle remains in the left-hand lane. After approximately 2-4 minutes, the congestion clears, and the surrounding traffic begins to accelerate, triggering a takeover request. Once the participant assumes manual control, they are expected to increase speed along with the other vehicles.

Participant reading a magazine. The display screen in the simulator shows "Automation Engaged", the simulated screen on the screen shows the road ahead.

Participant watching a film on a mobile phone in the driving simulator. The display screen shows "Take Control of Vehicle" with the simulated screen showing the road ahead.

Participant with both hands on the steering wheel and the display screen showing "Manual mode engaged" with the simulated screen showing the road ahead.

Driving simulation environment development

At UCL, the driving simulation environment was developed using the Unity game engine. Textures and 3D models for the simulated environment were sourced from the Unity Asset Store. At LU, scenarios were created using SCANeR studio v1.9, a sophisticated driving simulation software developed by AVSimulation. This software allows for the design of custom terrain environments, vehicle models, and scenario parameters. The SCANeR VISUAL module generates highly realistic synthetic images from the 3D environment, providing full immersion for drivers across three screens.

The simulation environments at both locations were matched as closely as possible given the different software and used the terrain of a generic UK motorway with three lanes of traffic in each direction, separated by a central barrier. Each scenario had three variations, differing only in the time taken to reach the point of the takeover request. In both scenarios, the ego vehicle (the car the participant is driving) began in automated driving mode and transitioned to manual mode upon the participant's input. The automated vehicle adhered to all road laws, drove at a fixed speed based on scenario parameters, and maintained its lane position, with no vehicles overtaking or changing lanes.

The ego vehicle, which the participant controls as they would a normal car, is equipped with autonomous capabilities, allowing it to merge into the slowest lane, follow the road layout, and maintain a constant pre-set speed or a safe distance from the vehicle in front. In manual mode, participants take control of both the lateral and longitudinal aspects of driving using the pedals and steering wheel. All other vehicles within the simulation are fully autonomous, programmed to stay within their lanes and follow the road layout.

Takeover request design

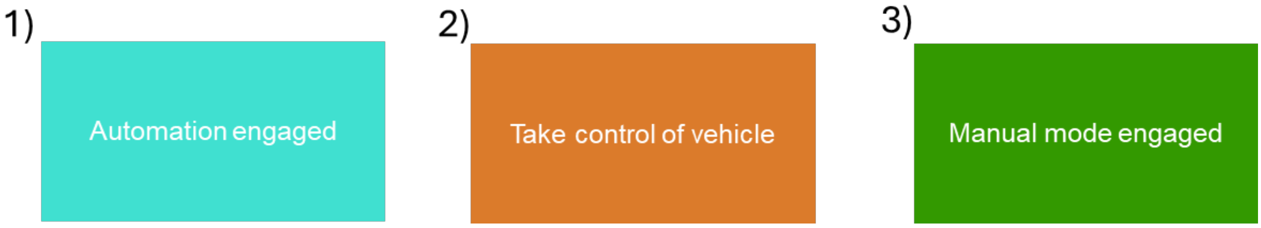


Figure 1 Stimulus shown to participants in the driving simulator to indicate a takeover request

UN ALKS regulations (2021) mandate that the status of the ALKS must be clearly indicated to the driver through a dedicated optical signal, typically within the driver's peripheral vision, such as a steering control or vehicle icon displaying an "A" or "AUTO". In this study, the takeover request was signalled by an auditory beep and direct written instructions ("Automation engaged", "Take control of vehicle", "Manual mode engaged") (see Figure 1). Written instructions were utilised in this study instead of symbols because it was important to eliminate the need for additional training on interpreting symbols. The written instructions provided clear, unambiguous guidance, ensuring participants could understand the required actions without prior familiarisation with the system.

The colour scheme and auditory alert design were developed based on current trends in ALKS enabled vehicles, where turquoise has been used to signify automation mode. For example, Mercedes-Benz incorporates similar visual and auditory cues in their S-Class vehicles, where an internal warning tone accompanies the takeover request.

At UCL, to take back manual control, the driver must either press one of the pedals or turn the steering wheel by more than 5 degrees. Upon doing so, they immediately regain manual control, and the screen turns green and reads "Manual mode engaged". In contrast, at LU, participants must verbally say "Ready to drive" and a researcher transferred control to them by pressing a button, after which the screen shows the green screen and reads "Manual mode engaged".

2.4 Procedure

The session began with a welcome and pre-clinic questionnaire, where participants provided baseline information relating to their attitudes to technology and transport innovations and signed a consent form. Following this, participants engaged in a practice drive in the simulator to familiarise themselves with the simulator controls and driving environment. Eye tracking glasses (and EEG at UCL) were fitted to the participant and an explanation of simulator sickness and monitoring of baseline symptoms was provided with reassurance that they should stop participating if any symptoms were experienced.

In each trial, participants received instructions on how to carry out the NDRA (or were informed there would be no NDRA). They were told that the system would be in self-driving mode but that they would receive a takeover request, at which point they were instructed to take manual control, **"as soon as you feel ready and safe to do so."** At UCL, manual control was given as soon as the participant made input to steering wheel or pedals. At LU, participants were required to say, “Ready to drive” and a researcher gave manual control.

The automated drives lasted between 2-4 minutes, each ending with a takeover request. Participants had up to 30 seconds to assume manual control and then drove manually for 30 seconds. Afterward, the simulation was paused, and participants completed the NASA-TLX and SART questionnaires. This process constituted one full trial. Participants performed eight different trials (six NDRAs and two No NDRA trials) across two scenarios. After the fourth trial, the simulator sickness questionnaire was administered again to monitor any changes, and participants were reminded that they could stop at any point.

After eight trials, participants left the driving simulator and were shown to an adjacent area where they completed a post-experiment questionnaire asking about their experience of the study. They were thanked for their time and paid an incentive for participation. Figure 2 shows a visualisation of the full experimental procedure.

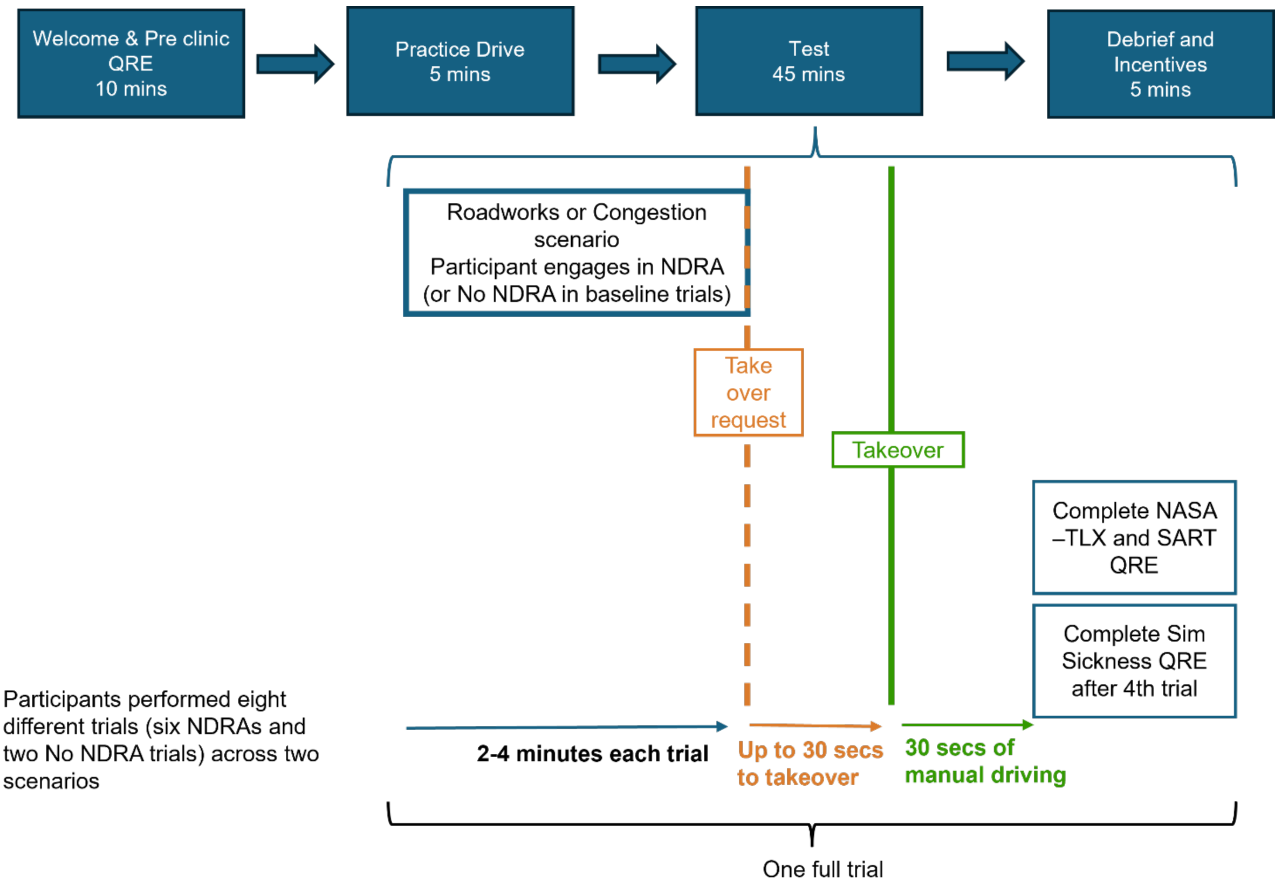


Figure 2: Overview of the research procedure.

Matrix sequencing

The order in which participants completed scenarios and NDRAs was rotated through six different sequences to prevent practice effects, and participants were randomly assigned to one of these sequences. The order within each matrix ensured that no NDRA activity was repeated in the same position across different matrices. For example, as shown in the Table 1 below, in Matrix 1, the drinking water activity was the first NDRA activity following the baseline activity. In the other five matrices, the drinking water activity never appeared as the first NDRA activity after the baseline activity. The matrices were also designed to evenly distribute the NDDTs across the two scenarios (i.e., for each NDDT, it was assigned to the Roadworks scenario for half the matrices, and the Congestion scenario for the other half).

The Roadworks and Congestion scenarios had 3 variations of differing lengths: 2mins 10sec, 3 mins, 3mins 50 sec, 2mins 20secs, 3mins 10 secs, 4mins. These variations were rotated amongst the tasks as well.

This design was implemented to ensure the fairness of the data, as having a single order could result in inaccurate data for the first (or last) NDRA activity due to participants still acclimating to the activities and understanding what was expected of them or having had more practice.

Table 1: Table shows the six matrices in a sequenced order in which participants performed the NDRAs and baseline ‘No NDRA’ conditions during two driving scenarios: Roadworks and Congestion.

Matrix 1	Matrix 2	Matrix 3	Matrix 4	Matrix 5	Matrix 6
Drinking water	Cradled mobile	Magazine	Eating popcorn	Handheld mobile	Wordsearch
Handheld mobile	‘No NDRA’ Roadworks	Eating popcorn	Drinking water	Cradled mobile	‘No NDRA’ Congestion
Magazine	Drinking water	Cradled mobile	Wordsearch	‘No NDRA’ Congestion	Eating popcorn
‘No NDRA’ Roadworks	Handheld mobile	Wordsearch	Magazine	Drinking water	Cradled mobile
Eating popcorn	Wordsearch	‘No NDRA’ Congestion	Handheld mobile	‘No NDRA’ Roadworks	Magazine
Cradled mobile	Eating popcorn	Drinking water	‘No NDRA’ Congestion	Magazine	‘No NDRA’ Roadworks
‘No NDRA’ Congestion	Magazine	‘No NDRA’ Roadworks	Cradled mobile	Wordsearch	Handheld mobile
Wordsearch	‘No NDRA’ Congestion	Handheld mobile	‘No NDRA’ Roadworks	Eating popcorn	Drinking water

Design of Non-Driving Related Activities (NDRAs)

We carefully considered the different types of activities explored in previous research and selected NDRAs that could be categorised into three distinct groups: Mobile phone use (requiring interaction with a mobile device), non-technological activities (more traditional and do not involve the use of digital devices), and motoric activities (physical interaction with objects). Two trials were categorised as “No NDRA” meaning that participants sat in the driver’s seat and did not engage in any activity.

For tasks involving technology, we chose not to use the built-in Human-Machine Interface (HMI) system since ALKS regulations dictate that the HMI would automatically stop upon a takeover request (United Nations, 2021). Instead, this study focuses on natural participant behaviour with the NDRA to observe how they chose to disengage from these tasks in preparation for manual control.

Mobile phone activities

- **Watching a film on a cradled mobile phone:** Participants were given the choice of a 5-minute YouTube video to watch on a Google Pixel 6a smartphone, which was securely attached to the dashboard to the left of the steering wheel and HMI. Options included a horror film ("Don't Look Away"), an animation ("Dustin"), a nature documentary ("Planet Earth"), or a TED talk ("The Power of Negative Thinking"). The researcher set up the selected video in the cradle, and participants started and paused the film using the play and pause buttons once the trial began.

- **Holding a mobile phone and playing a game:** Participants played Tetris on a Google Pixel 6a smartphone. Tetris is a classic puzzle game where players fit falling blocks into a grid to form complete horizontal lines, which then clear from the screen. The game ends when the stack of blocks reaches the top, leaving no space for new blocks. The game includes a fast-paced audio soundtrack that increases in intensity as time progresses. It was chosen as it is a widely familiar immersive game that can be played on a handheld device. The mobile device was placed on the passenger seat, and participants were instructed to pick it up, press play while holding the device, and begin playing the game once the trial started and automation was engaged (Jiang et al., 2024).

Non-technological activities

- **Reading a magazine**: Participants were offered a choice among four magazines: BBC Top Gear (cars), National Geographic: Traveller (travel), Take a Break (gossip), or BBC Good Food (food/cookery). The selected magazine was placed on the passenger seat, and participants were instructed to pick it up and start reading once the trial began and automation was engaged.

- **Completing a wordsearch or Sudoku**: Participants could choose between a word search puzzle book (Puzzler) or a Sudoku book (Sudoku Selection), both of which came with a pen. The puzzle book was placed on the passenger seat, and participants were instructed to start solving it once the trial started and automation was engaged.

Motoric activities

- **Drinking a cup of water:** Participants were asked to drink water from a disposable, recyclable coffee cup to simulate drinking during the automated drive. The cup was securely fitted with a lid and placed in the cup holder beside the participant. They were instructed to take frequent small sips of water throughout the drive, ensuring no spills, and to securely place the cup back in the holder after each sip to maintain safety. The cup was placed in the cup holder by the researcher, and participants were instructed to start taking sips once the trial started and automation was engaged.

- **Simulated eating of "Popcorn"**: To mimic the action of eating without the risks associated with food, participants were fitted with a cup holder on their chest just underneath their chin to act as a "mouth". They were provided with a packet of cotton balls. Throughout the automated drive, participants were instructed to transfer the cotton balls from the packet into the holster cup on their chest, simulating the action of eating a snack like popcorn (taking them out of the bag and lifting them to near their mouth and then dropping them into the cup holder). This approach avoided potential issues like allergies, messiness, or greasy hands on the steering wheel, which could affect the takeover process or the experience of subsequent participants.

  The cotton balls were placed on the passenger seat and participants were instructed to start 'eating' once the trial started and automation was engaged.

All NDRAs were thoroughly explained to participants before the trial began, and they were given the opportunity to ask any questions. However, no specific instructions were provided regarding what to do with the NDRA when the takeover request was issued. The only instruction given was to take over control "as soon as they felt ready and safe to do so."

Visual of the participant inside the simulator with a bag of cotton balls completing the "eating popcorn" NDRA and the display screen is on "Automation Engaged" and the simulated scene on the screen showing the road ahead.

2.5 Survey data

Before the experiment, participants completed a questionnaire designed to assess their familiarity with various technologies and their attitudes toward non-driving related activities (NDRAs) in non-driving tech. This survey included questions about:

- Ownership of items such as smartphones, Amazon Echo devices, wireless cameras, and VR headsets.
- Driving behaviours, including frequency, purposes, and habits.

- Comfort level with future transport innovations like ALKS, electric bikes, autonomous vehicles, and parking assist systems.
- Attitudes towards engaging in NDRAs while using non-driving technologies.

This pre-experiment data provided insights into participants' perspectives on self-driving vehicles, their level of immersion in technology, and their opinions on which activities should be permissible during self-driving.

After the experiment, participants answered additional questions aimed at evaluating their comfort, immersion, and understanding of the activities. Questions included, "*Were there any aspects of the experiment that you found particularly enjoyable or challenging*?" and "*How did you perceive the takeover request alerts during the experiment? Were they clear and easy to understand*?" These questions also explored whether participants' views on NDRAs and their immersion levels were influenced by the experiment. Full copies of the questionnaires are in Appendix 5.4.

2.6 Analysis

A total of 96 participants were initially recruited for this project. After excluding data from nine participants due to non-compliance with instructions, simulator sickness, or missing eye-tracking data, 87 participants' data were available for analysis. Data from UCL and LU is analysed separately due to differences in testing locations and unbalanced sample sizes, with the results presented here side by side for comparison.

Statistical differences between NDRAs and scenarios were assessed using Linear Mixed-Effects Models (LMM), which account for within-participant variability. Fixed effects were included for NDRAs and Scenario, while participant and location were treated as nested random variables. Further details on these standardised tests are provided in Technical Appendix 7. Statistical significance was set at $p < 0.05$.

To visualise the data across NDRAs, scenarios, and locations, both bar charts and raincloud plots are used. Raincloud plots are particularly effective in displaying comprehensive data by combining multiple chart types, providing a clear view of overall trends and individual data points. A tutorial on how to interpret these plots is available in Appendix 5.4.

2.7 Rationale of Situational Awareness metrics used in this study

A multifaceted approach that incorporated various metrics and methods was utilised to measure SA in this study. These measures were selected in line with the research discussed in Section 1, to capture both the behavioural and cognitive aspects of how participants responded to the transition from automation to manual driving. Figure 3 tracks the process from when a takeover request is issued to evaluating participants' reactions and perceptions during a driving scenario.

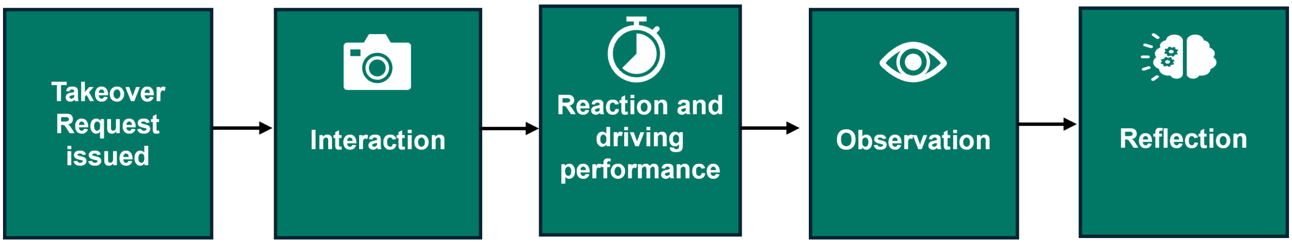


Figure 3: Stages of Takeover Process and Situational Awareness Assessment

The analysis was designed to investigate the following Key Performance Indicators (KPIs): Interaction, Reaction and driving performance, Observation and Reflection. These measured whether participants fully disengaged from NDRAs, how long they took to take over and whether participants took appropriate behavioural actions, such as adjusting speed based on the scenario, their visual attention patterns post-takeover and whether they perceived they had had good SA after the trial was over.

The metrics used to measure SA in each of these KPIs were as follows:

Interaction

We used GoPro cameras installed inside the vehicle to record participants' engagement and disengagement with Non-Driving Related Tasks (NDRTs). This allowed us to assess how frequently and for how long participants interacted with tasks unrelated to driving, providing insights into the level of immersion each activity produced, which would impact building SA after the takeover request.

Engagement with NDRAs:
Recording began as soon as participants picked up an NDRA, with researchers manually documenting the interactions, which were cross-checked with eye-tracking data. These observations were free form, aiming to understand how participants engaged with NDRAs during driving scenarios without using a predefined coding frame. How participants initiated and managed the NDRA, such as picking up a magazine, eating popcorn, or using a mobile phone, and whether they used one or both hands was recorded. Non-verbal behaviours, such as fidgeting with the NDRA or glancing at the dashboard, were also recorded to assess the level of engagement and distraction.

Disengagement from Non-Driving Related Activities (NDRAs):
Using recordings from in-cab GoPro cameras, whether participants fully or partially disengaged from the NDRAs was assessed. This was measured in a binary manner—either disengagement occurred, or it did not. "No NDRA" trials were excluded from this metric since there was no activity to disengage from. Appendix 5.1 outlines the coding determination of a full or partial disengagement in the context of each NDRA.

Reaction and performance

Reaction time was a critical measure, capturing how quickly participants took control of the vehicle after a takeover request. This metric is essential for understanding drivers' readiness and responsiveness in different scenarios. Performance indicators of driving once manual control was resumed was also evaluated.

Time to Takeover (TtTO):
Takeover time was defined as the time between the issuance of the takeover request and either the participant's first manual control input (e.g., steering or braking) at UCL, or the time it took them to respond “Ready to drive” at LU. Although ALKS regulations typically require a User-in-Charge to take control within 10 seconds of a takeover request, in this study, participants were given up to 30 seconds to respond. This extended time frame was set as the limit before reaching a safety-critical point in the scenario, beyond which failing to take over would be considered a failed attempt (as the ego car would have reached the roadworks). By allowing up to 30 seconds, how many participants responded within the 10-second window was analysed and captured a broader range of reactions, including those from participants who took longer to respond.

Time to Target Speed (TtTS):
Time to Target Speed (TtTS) was measured to assess how quickly participants responded to cues in the driving environment after taking over control from an automated system. This measure is closely related to their SA because it reflects their ability to perceive, understand, and react to changes in their surroundings, such as road signs in the roadworks scenario or other vehicles speeding up in the congestion scenario.

Time to Target Speed was evaluated differently in each scenario as follows:

- Roadworks Scenario: Time delay from when the participant crossed 75 metres before the gantry (where they can see the 50mph speed limit sign) until they reached the target speed of 60mph.
- Congestion Scenario: Time delay from the first input to either the steering wheel or pedals until reaching the target speed of 45mph.

If the target speed was not reached within 30 seconds, the time was recorded as NA, indicating that the participant did not significantly decelerate (roadworks) or accelerate (congestion).

Lane deviations:
Swerving suggests that participants may have taken over control without fully regaining SA, leading to difficulties in managing the vehicle, increasing the likelihood of lane deviations. Lane deviations were evaluated in two ways: through simulator data indicating diversions greater than half a car width, and by visual verification of the wheels crossing the white line without entering the adjacent lane (excluding deliberate lane changes).

Observation

We assessed how participants scanned their driving environment to build SA. Eye-tracking technology helped us understand where and how often participants looked at different areas, such as mirrors and road signs. The primary analysis included all AOIs to provide a comprehensive view of where participants directed their attention during this transition. A secondary analysis was conducted to specifically examine participants' use of mirrors across different NDRAs. The rationale for this focused analysis is grounded in the importance of mirrors for monitoring the vehicle's surroundings, especially during the transition from automated to manual driving, when drivers need to quickly assess their environment to make safe driving decisions. Physiological changes in pupil size and brain activity were also recorded to monitor shifts in SA, particularly in how participants observed their surroundings and how this awareness changed after the takeover request.

Areas of Interest (AOIs):
Where participants looked in the driving environment was analysed and for how long, directly after the takeover request and during the 30 seconds of manual driving. These Areas of Interest (AOIs) included:

- "HMI" (Human-Machine Interface): This AOI includes any interaction with in-vehicle systems, such as infotainment screens or control panels. Monitoring this area helps assess how participants' attention shifts to or from the vehicle's control systems, which could influence their ability to regain full control of the vehicle.
- "NDRA" (Non-Driving Related Activities): This AOI encompasses visual engagement with NDRA including glances back to the NDRA while manual driving which could indicate visual distraction.
- "Other Areas": This AOI captures instances where participants looked outside the predefined AOIs, potentially indicating searching behaviours around the vehicle or outside the cabin. Monitoring these glances can provide insights into how participants scan their environment for additional information that could affect their driving decisions.
- "Rear-view Mirror": The rear-view mirror is a key element in maintaining SA, allowing the driver to monitor traffic and other potential hazards behind the vehicle. Tracking looking times to this AOI is essential for understanding how well participants are assessing their surroundings when taking over control.
- "Right-Side Mirror": Similar to the rear-view mirror, the right-side mirror is crucial for checking blind spots and ensuring safe lane changes. Attention to this AOI provides additional context on how participants manage their visual scanning and SA of the vehicle's immediate surroundings.
- "Road": Primary focus area directly ahead of the vehicle. This is where drivers typically direct their attention to maintain lane positioning and anticipate upcoming obstacles.
- "Speedometer": Observing participants' attention to this AOI provides insight into their awareness of the vehicle's speed, which is crucial for maintaining control after a takeover request and complying with the contextual needs of the takeover scenario.

The selected AOIs were chosen to encompass the key areas where drivers typically direct their attention during driving, each contributing to a different aspect of situational awareness and driving performance. Analysing looking times across these AOIs, provides a detailed understanding of how participants allocated their visual attention during the critical transition period. Full details of the analysis pipeline can be found in Appendix 5.7.

Electroencephalography (EEG):
At UCL, EEG data was collected to analyse drivers' brain responses to the takeover requests to measure cognitive load and SA during the transition from automation to manual control. This data is analysed in a separate report.

Pupil diameter change rate:
The rate of the change of the pupil diameter in millimetres per second was also measured. When taking over control from a self-driving car, changes in pupil diameter can be a useful physiological indicator of a driver's SA. An increased pupil diameter change rate typically signals heightened cognitive load or stress, which can occur when a driver is trying to regain full awareness of the driving environment after a period of automated driving. Conversely, a lower change rate might indicate that the driver is more relaxed or less engaged. The pupil diameter change was taken and divided by the duration between the onset of the takeover request alarm and the reaching of the target speed.

Reflection

We employed questionnaires to gather subjective assessments of SA and workload. Participants provided feedback on their awareness and perceived workload during NDRTs. This subjective data was crucial for understanding their internal states and how different tasks impacted their SA and cognitive load.

By combining these methods, we capture a holistic view of SA, encompassing both objective performance metrics and subjective experiences. This comprehensive approach allowed us to identify how different tasks and scenarios influence User-in-Charge's SA providing valuable insights for enhancing safety and performance in automated driving systems. Participants completed the NASA-TLX and SART questionnaires after each trial to assess their perceived workload and whether their self-reported SA aligned with the objective measures. Walker et al (2008) reasoned that the simulated environment in driving simulators suppress SA. Participants were asked how realistic they found the simulations and which NDRAs they feel were challenging. Workload questions were followed by SART questions, assessing their SA for their take over for the NDRA they had just completed.

- SART evaluated the participants' perception, comprehension, and projection regarding the trial they had just completed. Ratings are made using 7-point Likert scales, captured post-drive, where 1 is labelled 'low', and 5 labelled 'high' and totals summed.
- NASA-TLX: Participants completed the NASA-TLX questionnaire after each trial on a tablet situated on the passenger seat, providing ratings on various dimensions of workload such as mental demand, physical demand, and perceived performance. We

asked participants to self-report visual, physical, and temporal workload pressures, as well as feelings of stress and attentional demands via the NASA TLX-R Workload measures (Hart, 2006). Responses were quantified on a sliding scale where 'very low' was 0 and 'very high' was 100. We only used five of the seven dimensions, due to the similarity of attention and visual workload dimensions to questions on the SART questions, so not to confuse participants with repetition.

By combining these methods, a holistic view of SA, encompassing both objective performance metrics and subjective experiences was investigated. The following section outlines the results and analysis.

3. Results

3.1 Interaction: Engagement with NDRAs during automated driving

Key findings

Mobile Phone Activities:

- Younger participants were more likely to use one hand when playing Tetris, while older participants used both, impacting phone handling during driving.
- Many struggled to turn off games or films, causing distractions during manual driving.

Non-Technological Activities:

- Reading magazines led to frequent glances at the road, meaning that participants may have had partial SA even though they didn’t need to monitor the road.
- Completing puzzles slowed the takeover process as both hands were occupied (book and pen).

Motoric Activities:

- The physical task of holding the cup and returning it to the holder proved challenging for many.
- Some participants managed this activity one-handed, keeping one hand on the steering wheel, while others continued eating during manual driving.

To gain insights into how participants interacted with non-driving related activities (NDRAs) and whether they fully disengaged when prompted to take over control of the vehicle, GoPro cameras were strategically placed within the driving simulators to capture detailed footage of participants' behaviour. These cameras provided a clear view of the participants' actions during the trials, including how they managed their NDRAs at the moment of the takeover request.

The footage was manually coded by trained observers to categorise and assess participant behaviour, particularly focusing on whether participants fully or partially disengaged from their NDRA upon receiving the takeover request. To ensure the reliability of these observations, the manually coded data was cross-checked with eye-tracking data, allowing for a comprehensive analysis of both visual attention and physical interactions during the critical transition from automated to manual driving. This dual approach ensured a robust assessment of participant behaviour, providing valuable insights into the effectiveness and safety of the takeover process.

Mobile phone activities

When playing Tetris on the handheld mobile phone, younger participants were more likely to use the phone one-handed, while older individuals tended to use both hands. This variation in grip significantly influenced the likelihood of dropping the phone before assuming manual control. Many participants could not turn the game off properly, even when they had put it on the passenger seat and taken over control it continued to keep playing the music, which was distracting and many participants made frequent glances back to it or even attempted to turn it off again during manual driving.

Watching a film on a cradled mobile phone presented additional challenges. Genres such as the animated film and the horror film, and documentaries required visual attention due to the moving scenes, but the TED talk was primarily auditory, allowing for intermittent visual disengagement. Participants who selected the TED talk tended to look around and just listen to it meaning they had a maintained some SA even while carrying out the NDRA. Many participants did not pause the film while assuming manual control either taking over immediately on the takeover request or attempting to pause it and then giving up and taking manual control anyway. Some participants completely changed their body position, lying across the seat or putting their head in the crook of their arm on the window, indicating they were very relaxed but also had to reposition themselves to get ready to takeover.

Non-technological activities

Reading a magazine had varying impacts on participants depending on the type of content. Eye-tracking data revealed that some individuals focused more on the pictures and just flicked through the magazine rather than reading text, which significantly influenced their level of visual engagement and reduced overall immersion in the activity. Frequent glances at the driving environment were observed, suggesting a lack of trust in the automated system and difficulty in maintaining focus on the reading material under the self-driving periods.

Completing a puzzle required participants to use both hands—one to hold the pen and the other to hold the puzzle book. Some people rested the book on their lap and some rested it on the steering wheel. This dual-hand engagement slowed the takeover process, as participants needed to set both items aside before taking control of the vehicle. Some participants began driving while still holding the pen, indicating a partial disengagement from the driving task but others took longer to put the lid back on the pen (only at UCL as the pen used at LU was a click type) which could have led to minor delays in taking over.

Motoric activities

The physical task of holding the cup and returning it to the holder proved challenging for many. Some participants continued to hold the cup even after assuming manual control of the vehicle, but most participants were not holding the cup when the takeover request occurred meaning there was no need to disengage from the task. While they took frequent sips participants looked around and tended to look down when putting the cup back in the holder.

Similarly, eating popcorn presented its own set of difficulties. Some participants managed this activity one-handed, keeping one hand on the steering wheel, while others continued eating during manual driving. Participants also struggled with determining when to stop eating and how to manage the popcorn while taking control of the vehicle. Some participants dropped the bag into their lap and took over with it still there, which could be problematic when manually driving if they spilled out (remember these were only cotton balls but they could have been some messy or hot food).

3.2 Interaction: Disengagement of the NDRA following a takeover request

Key findings

- Mobile phone-related NDRAs resulted in the lowest disengagement, indicating participants often started manual driving without stopping the activity.
- Reading a “Magazine” had the highest disengagement rates at both UCL (95.7%) and LU (90.3%).
- Location had a significant effect on disengagement, with UCL participants being less likely to disengage compared to LU ($p < 0.001$).
- No significant difference in disengagement between Roadworks and Congestion scenarios ($p = 0.167$).

Full disengagement was recorded as completely stopping engaging with the NDRA and putting it away (returning it back onto the passenger seat). Binomial models were utilised to investigate the effect of scenario type, activity, and location on disengagement. Full model details can be found in Technical Appendix 6.1, 6.2 and 6.3.

Figure 4 and Table 2 show the disengagement rate by Activity and Location.

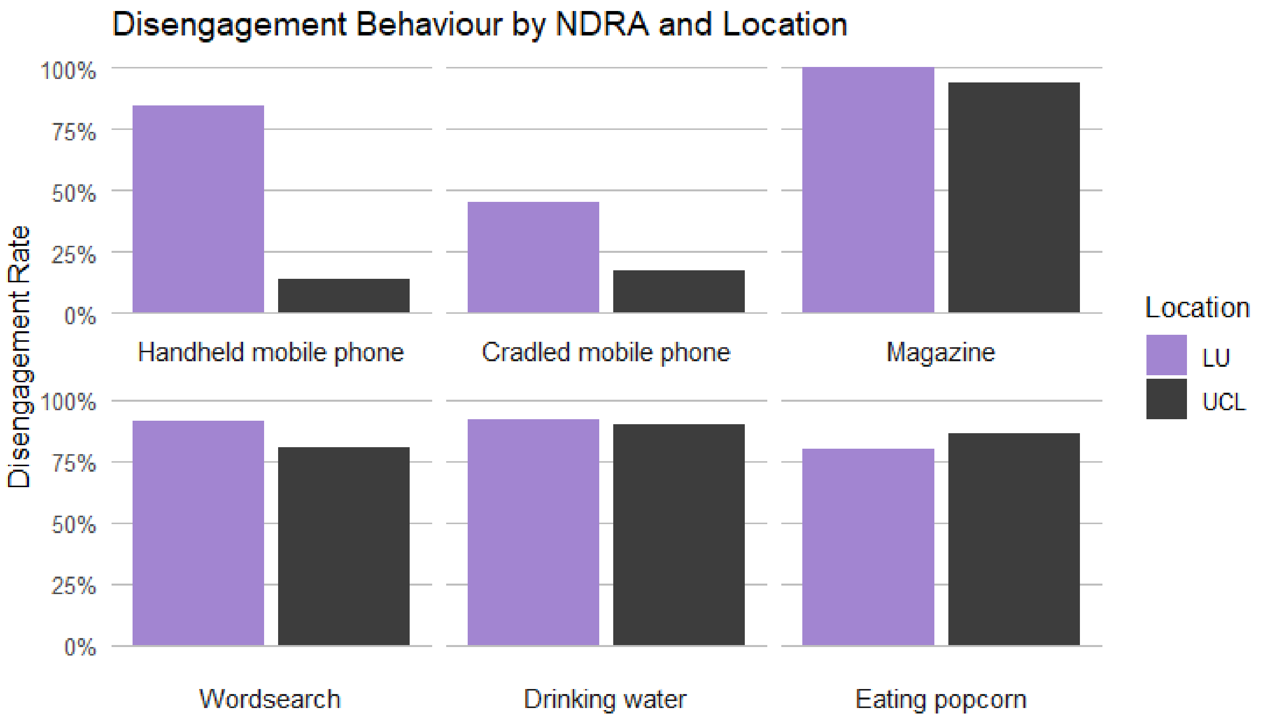


Figure 4: The bar chart titled "Disengagement Behaviour by NDRA and Location" compares the disengagement rates by NDRA and location.

Table 2: Table showing the percentage of participants who partially or fully disengaged from various activities at two locations, LU and UCL.

	LU		UCL	
Activity	**Partially Disengaged (%)**	**Fully Disengaged (%)**	**Partially Disengaged (%)**	**Fully Disengaged (%)**
Cradled mobile phone	56.5	43.5	84.5	15.5
Drinking water	12.5	87.5	11.9	88.1
Eating popcorn	20.8	79.2	18.6	81.4
Handheld mobile phone	17.4	82.6	87.9	12.1
Magazine	4.3	95.7	9.7	90.3
Wordsearch	16.7	83.3	20.3	79.7

Results showed significant variability in disengagement behaviour across the different NDRAs. When engaged in activities like "Drinking Water ($p < 0.001$)," "Eating Popcorn," ($p < 0.001$), doing a wordsearch ($p < 0.001$), and reading a "Magazine" ($p < 0.001$), participants showed high disengagement rates (above 75%) across both locations. Participants were significantly more likely to disengage from these NDRAs compared to using a handheld mobile phone ($p < 0.001$) or when using a cradled mobile phone ($p < 0.001$).

The “Cradled mobile phone” activity showed much lower disengagement rates at UCL (15.5%) compared to LU (43.5%). The lowest disengagement was observed for the “Handheld mobile phone” activity at UCL (12.1%), while at LU, it was for the cradled mobile phone. Despite the difference in rates, in both locations, NDRAs involving mobile phone use resulted in the lowest disengagement, suggesting that participants often resumed manual driving without pausing the film or completely stopping the Tetris game. Conversely, reading a magazine led to the highest disengagement at both locations, with 95.7% at UCL and 90.3% at LU.

Analysis also revealed a significant effect of location on disengagement behaviour, with participants at UCL being significantly less likely to disengage compared to those at LU ($p < 0$). However, it is important to note that the LU sample was much smaller ($n = 24$) than the UCL sample ($n=63$), which may affect the robustness of this finding.

There was no significant difference in disengagement behaviour between the Roadworks and Congestion scenarios ($p = 0.167$). This suggests that when controlling for the type of activity participants were engaged in, the scenario type did not significantly influence the likelihood of disengagement.

3.3 Reaction and performance: Time to Takeover by Activity and Location

Key findings

- Participants generally took longer to take over control for the handheld mobile phone NDRA, with noticeable differences in times between the two locations for the cradled mobile phone NDRA.
- Non technological activities like reading a magazine or doing a wordsearch had longer takeover times than other NDRAs.
- LU participants showed more variability in their response times compared to UCL.
- Physical tasks like drinking water or eating popcorn showed more variability in response times, especially at LU. UCL participants generally showed faster takeover times after these NDRAs.
- The absence of any distractions in No NDRA trials led to the fastest takeover times, reflecting the ease of response when no task disengagement was required.
- There was no significant effect of location on NDRA performance, although there was a significant effect observed in No NDRA trials, indicating that location influenced takeover times only when participants were not engaged in a task.

This section examines the time it took for participants to take control of the vehicle after a takeover request. The takeover was measured by input to either the steering wheel or pedals at UCL or by saying "Ready to drive" at LU analysed across two scenarios and eight trials. It's important to note that a shorter takeover time does not necessarily indicate better performance, as participants might not fully disengage from the NDRA, potentially leading to reduced SA and unsafe driving behaviour.

A linear mixed-effects model (LMM) was employed, with activity and scenario treated as fixed effects, and participant and location as nested random effects. Detailed information on the fixed and random effects can be found in Technical Appendix 6.4. The smaller sample size at LU (n=24) may have contributed to the increased variability observed in the data, which should be taken into account when interpreting these results.

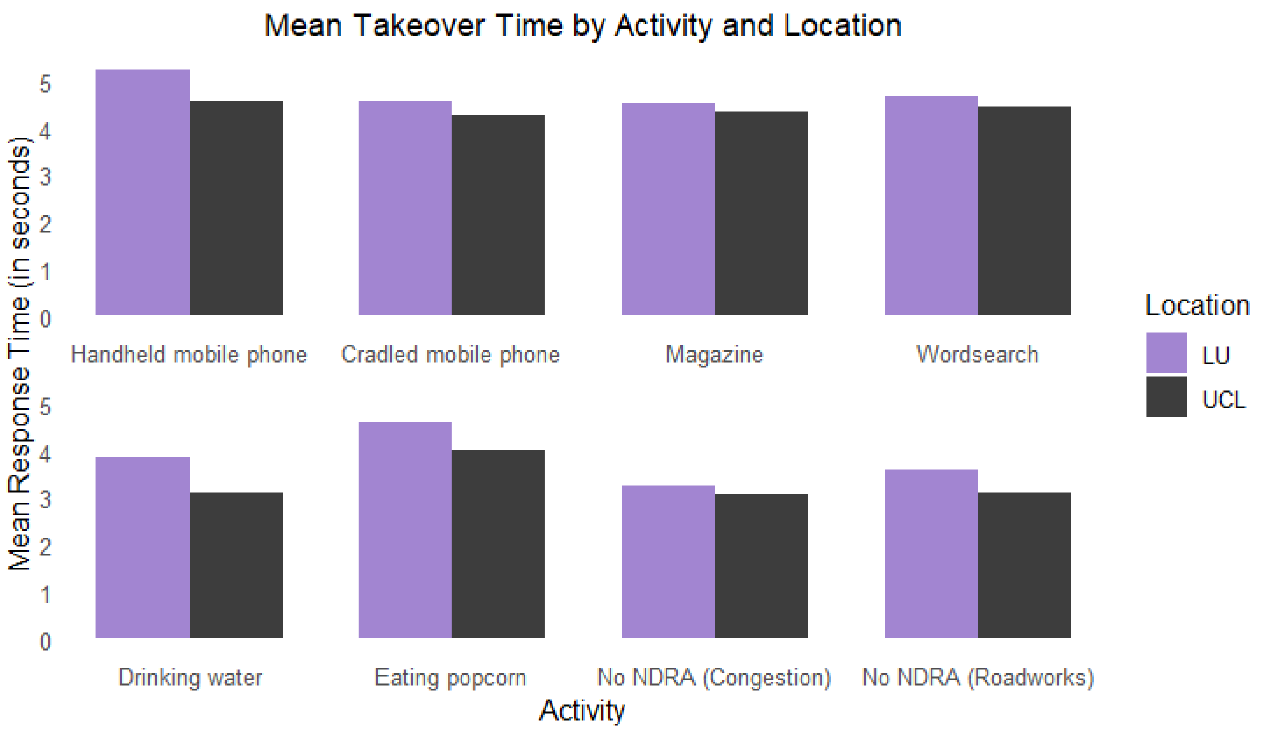


Figure 5: The bar chart titled "Mean Takeover Time (in secs) by Activity and Location" displays the mean response times for participants to take over manual control of a vehicle after engaging in various Non-Driving Related Activities (NDRAs).

Table 3: Table showing the means and standard deviations of Time to Takeover (TtTO) for various activities at two locations, LU and UCL. The table has four columns: Location, Mean Takeover Time in seconds and Standard Deviation in seconds. For each activity, the table lists the means and standard deviations in seconds at each location.

	LU		UCL	
Activity	**Mean Takeover time (seconds)**	**Standard deviation (seconds)**	**Mean Takeover time (seconds)**	**Standard deviation (seconds)**

Cradled mobile phone	4.57	2.45	4.28	3.43
Drinking water	3.87	1.92	3.11	1.91
Eating popcorn	4.59	2.51	4.02	3.12
Handheld mobile phone	5.25	3.51	4.58	3.18
Magazine	4.50	2.83	4.33	3.39
No NDRA (Congestion)	3.24	1.55	3.08	2.70
No NDRA (Roadworks)	3.58	1.95	3.12	2.12
Wordsearch	4.67	2.66	4.43	2.19

The comparison of the top three activities with the longest mean takeover times reveals some differences between the LU and UCL sites (Table 3 and Figure 5). At LU, participants took the longest time to take over when using a handheld mobile phone (5.25 seconds), followed by engaging in a wordsearch (4.67 seconds), and eating popcorn (4.59 seconds). In contrast, at UCL, the longest takeover times were also associated with using a handheld mobile phone (4.58 seconds) and engaging in a wordsearch (4.43 seconds), but the magazine activity was ranked third in terms of duration (4.33 seconds).

These differences suggest that the tasks themselves might have been perceived or interacted with differently across the two locations. For example, the mean time to take over control was generally around 3-5 seconds, with the "Handheld mobile" task taking the longest (LU: 5.25 seconds, UCL: 4.58 seconds). LU participants took slightly longer across all NDRAs, possibly due to the "Ready to Drive" protocol and the delay before manual control was given, although this difference was not statistically significant ($p = 0.16$).The shortest takeover times were observed in the No NDRA conditions (Congestion: LU 3.24 seconds, UCL 3.08 seconds; Roadworks: LU 3.08 seconds, UCL 3.12 seconds), and both were highly significant ($p < 0.001$), as expected since there was no activity to disengage from in those trials. The effect of scenario on time to takeover is discussed in more detail below.

The "Eating popcorn" task at LU (4.59 seconds) took longer than at UCL (4.02 seconds), although this difference was not significant ($p = 0.32$). "Drinking water" was the quickest NDRA for takeover time at both locations, with this difference being statistically significant ($p < 0.001$), likely because participants could still monitor the road and may not have been holding the cup at the time of the takeover request.

However, means do not capture the full range of participant performance, particularly the extremes where people perform very well or very poorly. Raincloud plots provide a more nuanced view (for a detailed explanation, see Appendix 5.3).

The raw data reveals considerable variation in how participants performed within each site, with significant differences across NDRAs. The quickest TtTO was 0.70 seconds (for “Drinking water”) and the longest was 24.40 seconds (“Magazine”). This variability indicates that while some participants were very quick to take over, others were much slower. Interestingly, many NDRAs show two or three distinct peaks in the density plots, suggesting that the data is multi-modal (some takeover times are more frequent than others with no average).

Mobile phone activities

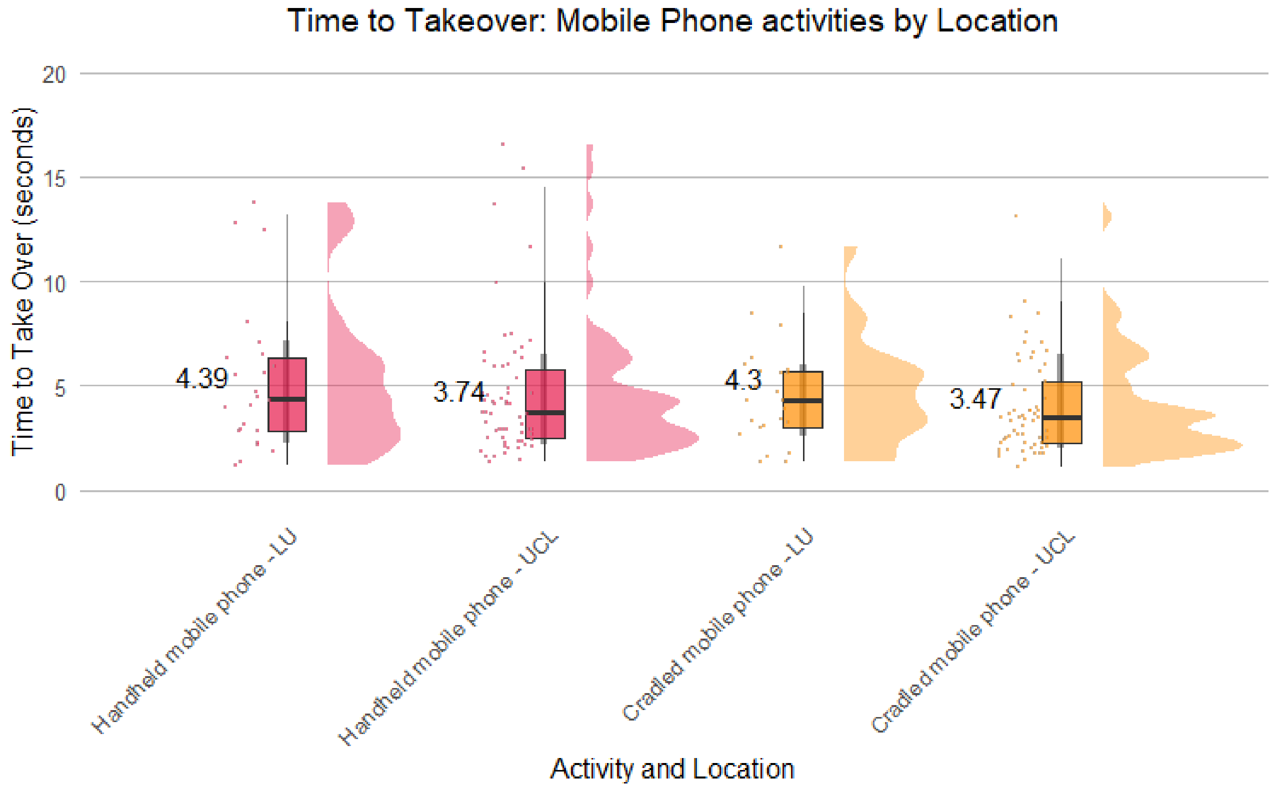


Figure 6: Raincloud plot displaying the time it took participants to take control of a vehicle after engaging in mobile phone activities, shown by location (LU and UCL).

Participants at the LU location had a median takeover time of 4.39 seconds when using a handheld mobile phone, showing variability in how quickly different participants responded (see Figure 6). At the UCL location, the median takeover time for handheld mobile phone use was slightly faster, at 3.74 seconds. The density curve suggests that most participants responded relatively quickly, with fewer outliers, although some took longer than 10 seconds to take over. However, the interaction between "Handheld mobile phone" and location was not significant ($p = 0.983$), indicating that the difference in takeover times between LU and UCL for this activity is likely due to random variation.

For cradled mobile phone use, the median takeover time at LU was 4.3 seconds, which was similar to the handheld condition at the same location. Participants at UCL had a quicker response time in this category, with a median of 3.47 seconds, a difference that was statistically significant ($p < 0.012$). This suggests that placing the phone in a cradle on the dashboard did not significantly reduce the takeover time compared to holding it in the hand at LU. However, the tighter distribution at UCL indicates that most participants there responded quickly and consistently, though a few still took longer. These longer response times tended to occur in the same participants across different NDRAs. The quicker takeover times for mobile phone activities may reflect the lower disengagement rates associated with these tasks, as discussed in section 4.3.1.

UCL participants were quicker to take over for the handheld mobile phone activity (3.74 seconds) and also had the lowest disengagement rates for that NDRA (15.5%), suggesting their speed was likely because many were not fully engaged in the task, allowing for quicker responses. The higher takeover time for LU (4.39 seconds) is likely due to almost half of the participants (43.5%) properly disengaging from the NDRA before resuming manual control. However, the differences in takeover times across NDRAs were not significantly different, and there was no significant effect of location ($p = 0.4051$), meaning the location of the test did not impact the takeover time. (See Technical Appendix 6.2 for full p-values for all tasks).

Non technological activities

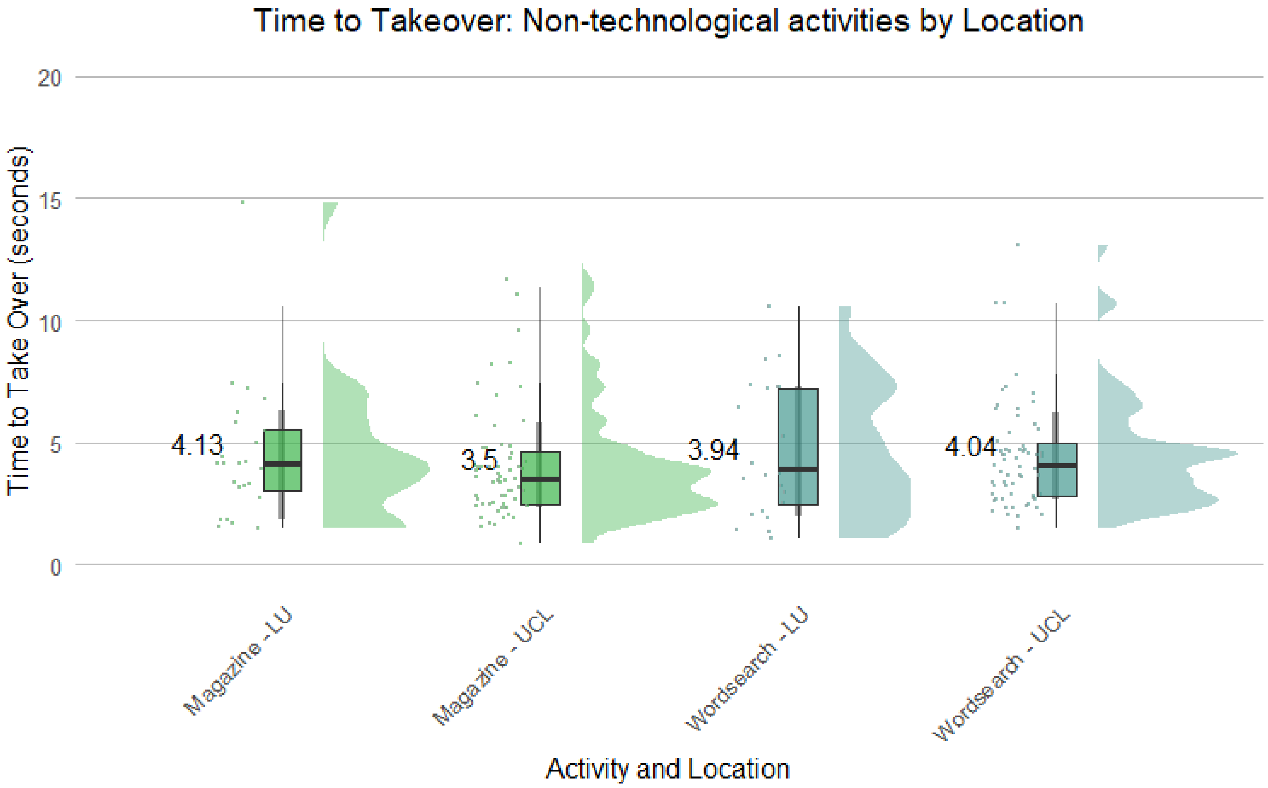


Figure 7: Raincloud plot displaying the time it took participants to take control of a vehicle after engaging in non-technological activities, shown by location (LU and UCL).

Participants at LU had a median takeover time of 4.13 seconds after reading a magazine. While there was some variability in the data, the overall distribution was relatively tight around this median value. At UCL, participants had a quicker median takeover time of 3.5 seconds. Although this was faster than the LU group, the difference was not statistically significant ($p = 0.798$), suggesting that UCL participants may have found it easier to disengage from the magazine task, but not to a significant extent (see Figure 7).

For the wordsearch NDRA, the median takeover time at LU was 3.94 seconds, slightly faster than for the magazine task. However, the data showed more variability, indicating that some participants took longer to take over. Again, this difference was not statistically significant ($p = 0.119$). UCL participants had a median takeover time of 4.04 seconds, which was comparable to the time taken for the magazine task at LU but slightly slower than their performance on the magazine task. Overall, the data suggests that UCL participants generally took less time to take over compared to those at LU, particularly for the magazine task.

Motoric activities

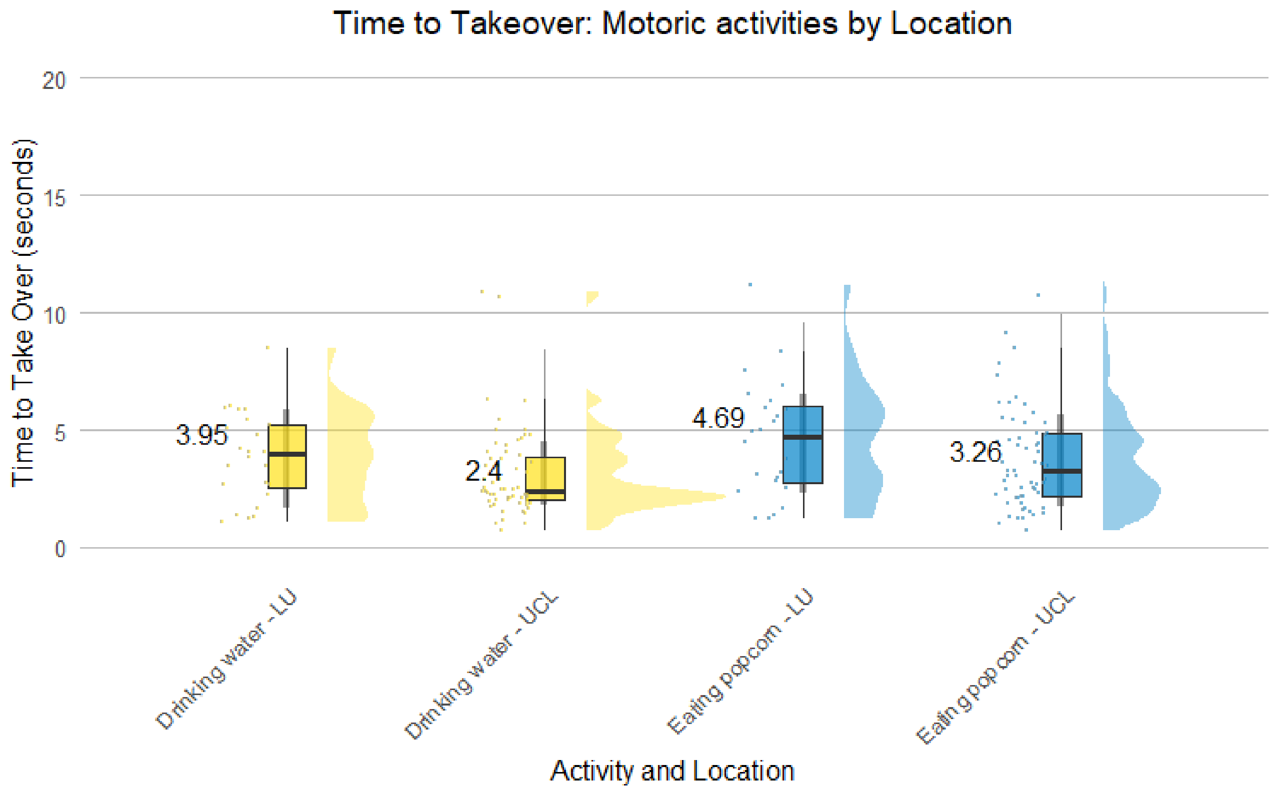


Figure 8: Raincloud plot displaying the time it took participants to take control of a vehicle after engaging in motoric activities, shown by location (LU and UCL).

Although participants at UCL appeared to take over control slightly faster (median of 2.4 seconds) compared to those at LU (median of 3.95 seconds) while drinking water, this difference was not statistically significant ($p = 0.4055$) (see Figure 8). This indicates that the difference in takeover times between locations for this activity is likely due to random variation rather than a meaningful effect of location. While the median takeover time at LU was higher, the variability within each group and the overlap in the distributions suggest that the locations did not meaningfully differ in how quickly participants resumed control.

Participants at LU took an average of 4.69 seconds to take over when 'Eating popcorn,' with a noticeably wider spread of data points compared to UCL, where the median was slightly lower at 3.26 seconds. This NDRA was also longer than the cradled mobile phone task (4.59 seconds). This difference in central tendency suggests that the physical task of “eating” may have caused more variability in response times at LU, possibly due to differences in task engagement or environment. The density curves in the raincloud plot indicate a slight positive skew in the 'Eating popcorn' activity at the UCL location, with a few participants taking substantially longer to regain control compared to the median. This suggests that while most participants were quick to resume driving, some experienced delays, possibly due to being more engrossed in the task, perhaps due its novel design (a holster round the neck and cotton wool balls as ‘popcorn’).

No NDRA

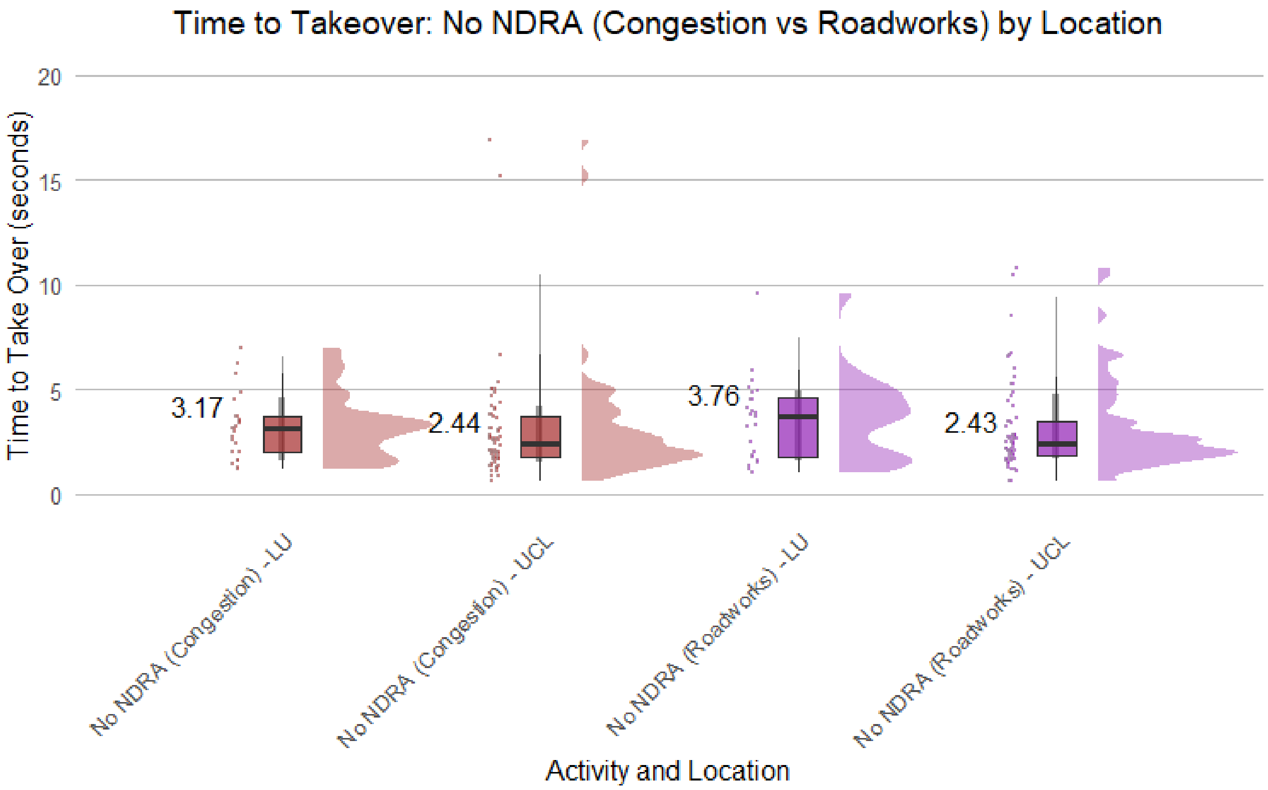


Figure 9: Raincloud plot displaying the time it took participants to take control of a vehicle when there was no NDRA, differentiated by congestion and roadworks, and shown by location (LU and UCL).

The "No NDRA" conditions, both in Congestion and Roadworks scenarios, resulted in relatively fast takeover times, which aligns with expectations since participants did not need to disengage from any task (see Figure 9). At LU, the median takeover time was 3.17 seconds in the Congestion scenario, and 3.76 seconds in the Roadworks scenario. The corresponding times at UCL were quicker, with medians of 2.44 seconds and 2.43 seconds, respectively. These differences were statistically significant ($p < 0.001$), indicating that participants at UCL were generally more alert and ready to take over when no distractions were present.

However, the slightly longer takeover times observed at LU, particularly in the Roadworks scenario, could suggest that environmental factors at LU or the perceived complexity of the roadworks required more cognitive processing time, despite the absence of an NDRA. This difference between scenarios was also significant ($p < 0.05$), suggesting that the scenario type itself plays a role in how quickly participants respond, yet the absolute differences in seconds are quite minor.

Interestingly, even in the absence of any NDRA, there was considerable variability in takeover times among participants. Some took over 5 seconds, with a few taking as long as 15 seconds. This variability could imply that these participants were building situational awareness (SA) before taking control, ensuring they had a complete understanding of the road environment before resuming manual control. However, despite these longer takeover times by some participants, the overall effect of location on takeover time in these no-task scenarios was not statistically significant (p = 0.4051), suggesting that location did not significantly influence the speed of the takeover in the absence of an NDRA.

3.4 Reaction and performance: Time to Takeover by Scenario and Location

Key findings

- Across both LU and UCL locations, participants generally took longer to take over in the Roadworks scenario compared to the Congestion scenario. This trend is consistent across most activities.
- Even in the absence of an NDRA, participants took slightly longer to take over in the Roadworks scenario, suggesting the scenario itself adds complexity.
- The comparison of takeover times between Congestion and Roadworks scenarios shows statistically significant differences, indicating the takeover scenario does influence how quickly participants can regain control but the difference in seconds is very small (less than half a second).

Scenarios like congestion and roadworks present different levels of complexity and urgency, potentially affecting how quickly a driver can regain control of the vehicle. Splitting the analysis by scenario can illustrate how these different takeover request contexts can impact takeover times.

A nested random intercept model was used to examine the effects of activity and scenario type on reaction times, with location and participant considered as nested random effects. In this model, activity and scenario type were treated as fixed effects, while participant was nested within location to account for the variability both between locations and between participants within each location. The smaller sample size at LU (n=24) may have contributed to the increased variability observed in the data, which should be taken into account when interpreting these results. An interaction model was also employed to test for an interaction between activity and location, meaning it examines whether the effect of the NDRA on reaction time differs depending on the location the participant carries out the study in. Further details about the models can be found in Technical Appendix 6.4 and 6.5.

The comparison of mean takeover times between the Congestion and Roadworks scenarios shown in Figure 10 reveals while both locations experienced an increase in variability during the Roadworks scenario, the mean takeover times did not differ drastically between the two scenarios or locations. However, LU participants consistently took slightly longer to take over control, and the Roadworks scenario introduced more variability in participants' responses. Tables 4 and 5 show the mean and standard deviations for takeover times by each scenario for each NDRA.

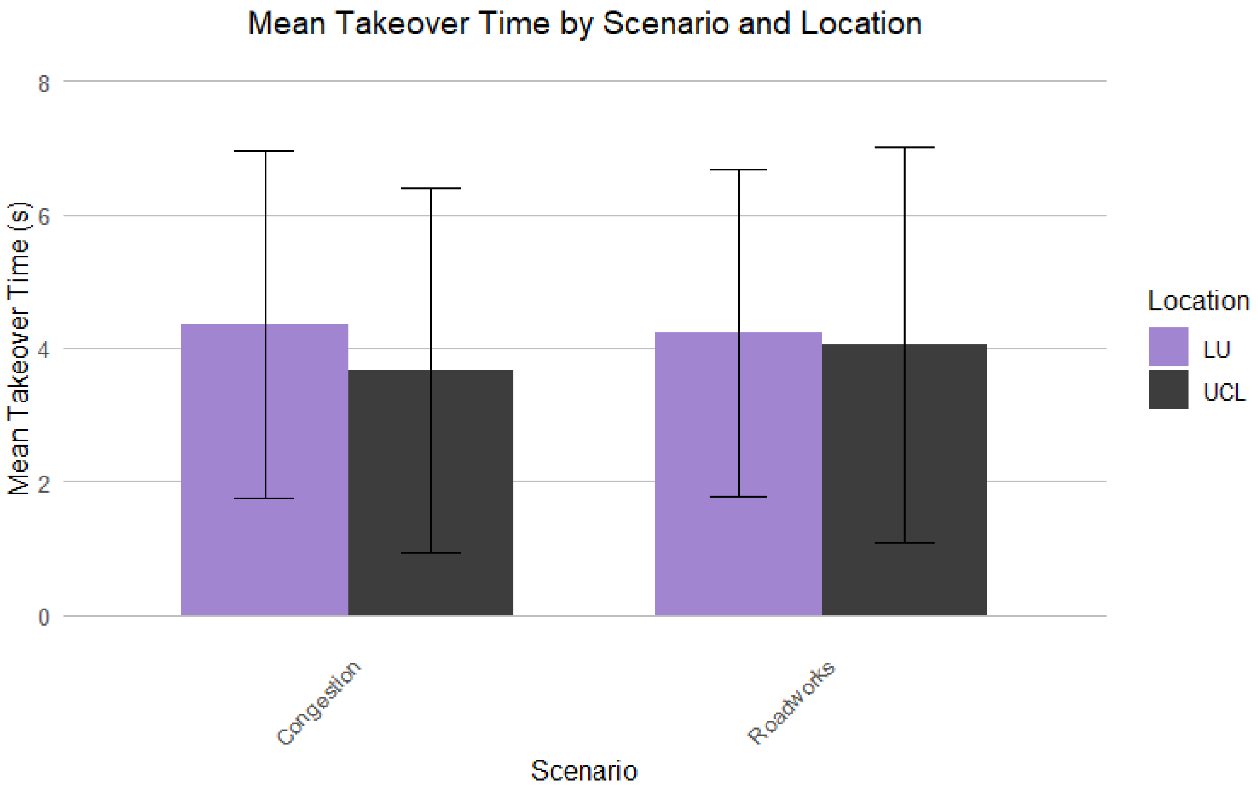


Figure 10: The image displays a bar chart which compares the mean takeover time in seconds for two scenarios, Congestion and Roadworks, across two locations (LU and UCL).

Table 4: Table showing the means and standard deviations of Time to Takeover (TtTO) for various activities during the Congestion scenario at two locations, LU and UCL. The table has four columns: Location, Activity, Mean Takeover Time in seconds, and Standard Deviation in seconds. For each activity, the table lists the means and standard deviations in seconds at each location.

	LU		UCL	
Activity	**Mean Takeover time**	**Standard Deviation**	**Mean Takeover time**	**Standard Deviation**
Cradled mobile phone	4.43	2.28	4.30	4.08
Drinking water	3.82	2.30	2.83	1.40
Eating popcorn	4.88	2.00	3.15	1.76
Handheld mobile phone	5.22	3.60	4.56	3.30
Magazine	4.69	3.60	4.05	2.44
No NDRA (Congestion)	3.24	1.55	3.08	2.70
Wordsearch	5.21	2.97	4.36	2.24

Table 5: Table showing the means and standard deviations of Time to Takeover (TtTO) for various activities during the Roadworks scenario at two locations, LU and UCL. The table has four columns: Location, Activity, Mean Takeover Time in seconds, and Standard Deviation in seconds. For each activity, the table lists the means and standard deviations in seconds at each location.

	LU		UCL	
Activity	**Mean Takeover time**	**Standard Deviation**	**Mean Takeover time**	**Standard Deviation**
Cradled mobile phone	4.71	2.70	4.26	2.79
Drinking water	3.91	1.62	3.37	2.28
Eating popcorn	4.18	3.16	4.94	3.93
Handheld mobile phone	5.28	3.59	4.59	3.10
Magazine	4.31	1.92	4.61	4.18
No NDRA (Roadworks)	3.58	1.95	3.12	2.12
Wordsearch	4.14	2.31	4.51	2.18

Across both scenarios, LU participants generally show longer mean takeover times compared to UCL participants. This trend remains consistent with previous discussions with some small variations:

- **Handheld mobile phone:** The mean takeover times slightly increased (LU: 5.22 seconds in Congestion to 5.28 seconds in Roadworks; UCL: 4.56 seconds in Congestion to 4.59 seconds in Roadworks). These differences were not statistically significant ($p = 0.169$).
- **Cradled mobile phone:** Both LU and UCL participants took longer in the Roadworks scenario (4.71 seconds and 4.26 seconds, respectively) compared to Congestion (4.43 seconds and 4.30 seconds, respectively). The effect of scenario on takeover time for this activity was statistically significant ($p = 0.046$).
- **Drinking water:** At LU, the mean takeover time increased slightly from Congestion (3.82 seconds) to Roadworks (3.91 seconds). UCL participants showed a more noticeable increase from 2.83 seconds in Congestion to 3.37 seconds in Roadworks. This increase in takeover time by scenario was statistically significant ($p < 0.001$).
- **Eating popcorn:** LU participants took less time in the Roadworks scenario (4.18 seconds) compared to Congestion (4.88 seconds). However, UCL participants showed an increase in takeover time in Roadworks (4.94 seconds) compared to Congestion (3.15 seconds). There was no statistically significant effect of the "Eating Popcorn" activity across scenarios ($p = 0.323$).
- **No NDRA:** The mean takeover times for the No NDRA condition were relatively consistent across scenarios. LU participants took slightly longer in the Roadworks scenario (3.58 seconds) compared to Congestion (3.24 seconds). UCL participants also took slightly longer in Roadworks (3.12 seconds) compared to Congestion

(3.08 seconds). The No NDRA conditions were statistically significant in both scenarios ($p < 0.001$ in the original model).

The comparison between Congestion and Roadworks scenarios reveals statistically significant differences in takeover times ($p < 0.001$). However, the variability observed in the LU data, potentially due to the smaller sample size (n=24), may have contributed to more pronounced fluctuations in the data distribution. While the Roadworks scenario generally leads to longer and more variable takeover times—particularly for activities like eating popcorn and drinking water—the actual difference in time is minimal (less than half a second). Although these differences suggest that the Roadworks scenario imposes a slightly higher cognitive load, they do not fundamentally change the overall conclusions, indicating that participants' ability to regain control is only modestly affected by the type of scenario.

3.5 Reaction and performance: Time to Target Speed (TtTS)

Key findings

- UCL participants took longer and showed more variability in reaching target speed, especially in the Roadworks scenario.
- Activities like "Cradled mobile phone," "Eating popcorn," "Drinking water," and "Wordsearch" significantly delayed time to target speed at UCL compared to LU.
- In the Congestion scenario, using a cradled or handheld mobile phone significantly increased time to target speed but NDRAs such as "Eating popcorn," "Magazine," and "Wordsearch" did not significantly alter time to target speed in Congestion.

In the Roadworks scenario, participants who had gained Level 2 SA (comprehension) by looking at one of the 50mph road signs should take appropriate dynamic driving action (start slowing down to appropriate speed). In the Congestion scenario, there was no explicit information in the environment in this scenario instructing participants on what to do. However, by observing other cars speeding up, participants should adjust their speed accordingly to match the speed of the surrounding traffic indicating some level of SA had been reached after the takeover request. A small proportion of participants failed to reach the target speed, indicating that they did not adequately adapt to the road conditions as required by the scenario. Specifically, 16.6% of UCL participants and 10.4% of LU participants did not reach the target speed.

Two linear mixed-effects models were used to analyse the time to target speed in different driving scenarios: congestion and roadworks. The models measured the effects of activity and whether participants reached target speed, with location and participant as random effects. Results are detailed in Technical Appendix 6.6, 6.7 and 6.8.

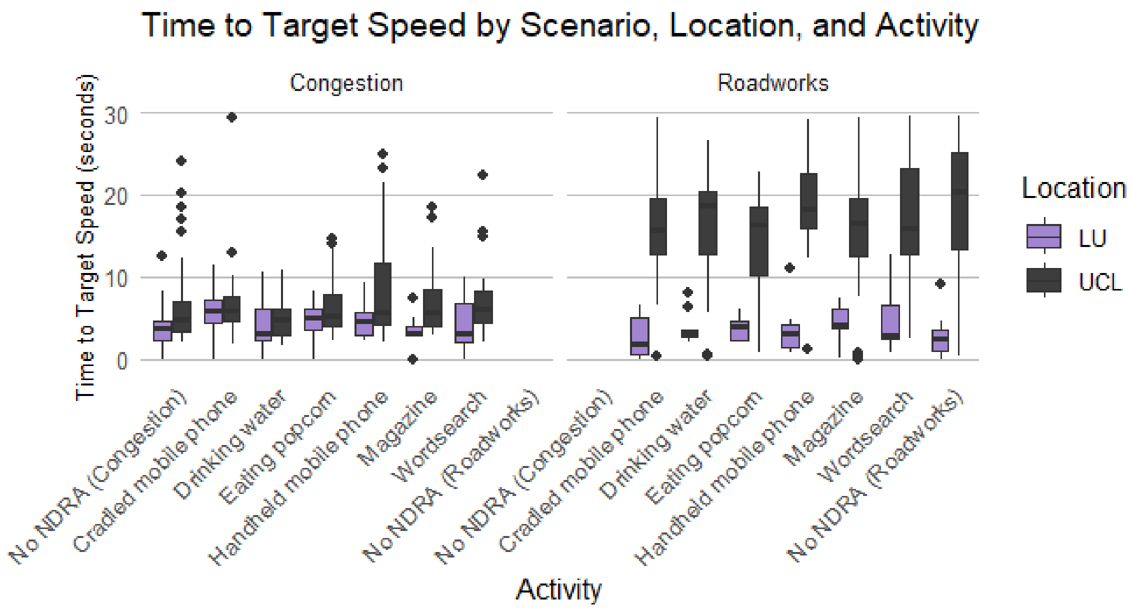


Figure 11: The image is a box plot titled "Time to Target Speed by Scenario, Location, and Activity." It displays the distribution of time (in seconds) it took participants to reach the target speed after a takeover request.

Table 6: The table presents the mean time to target speed and standard deviation, organised by activity, scenario type (Congestion and Roadworks), and location (LU and UCL). Each activity, such as Cradled mobile phone and Drinking water, is grouped under its respective scenario, with four columns showing the mean time to target speed and standard deviation for both LU and UCL.

		LU		UCL	
Activity	**Scenario**	**Mean time to target speed**	**Standard deviation**	**Mean time to target speed**	**Standard deviation**
Cradled mobile phone	Congestion	6.2	3.4	7.0	5.0
Drinking water	Congestion	4.3	3.2	4.9	2.4
Eating popcorn	Congestion	4.6	2.5	6.2	3.4

Handheld mobile phone	Congestion	4.9	2.2	8.7	6.3
Magazine	Congestion	3.6	1.9	7.1	4.1
No NDRA (Congestion)	Congestion	3.9	2.9	6.2	4.6
Wordsearch	Congestion	4.0	3.2	7.1	4.2
Cradled mobile phone	Roadworks	2.7	2.5	15.9	6.1
Drinking water	Roadworks	3.8	2.0	16.6	7.3
Eating popcorn	Roadworks	3.9	1.7	14.4	6.7
Handheld mobile phone	Roadworks	3.6	3.0	19.0	5.8
Magazine	Roadworks	4.6	2.0	15.5	8.3
No NDRA (Roadworks)	Roadworks	2.6	2.1	18.9	7.9
Wordsearch	Roadworks	4.7	3.6	17.3	7.1

UCL participants generally took longer and exhibited greater variability in the Roadworks scenario across all NDRAs (see Figure 11 and Table 6). For the "Cradled mobile phone" NDRA, UCL participants had a significantly longer mean Time to Target Speed (TtTS) of 15.9 seconds compared to 2.7 seconds at LU ($p = 0.01$). The "Eating popcorn" activity also resulted in significantly longer times, with UCL participants averaging 14.4 seconds versus 3.9 seconds at LU ($p < 0.001$). When reading a magazine, UCL participants took an average of 15.5 seconds to reach target speed, compared to 4.6 seconds at LU ($p < 0.001$). For "Drinking water," UCL participants needed 16.6 seconds on average, while LU participants took only 3.8 seconds ($p = 0.018$). The "Wordsearch" activity led to a mean TtTS of 17.3 seconds at UCL, sharply contrasting with 4.7 seconds at LU ($p <0.001$). TtTS for the "No NDRA" condition during the Roadworks scenario was not significant ($p < 0.34$). This indicates that when participants were not distracted by an NDRA, their TtTS was more consistent and less influenced by external factors, although there was still some variability.

The high standard deviations at UCL indicate significant variability among participants, but since the analysis was conducted within participants (comparing each participant against their own performance), these differences were not simply due to variations in driving ability.

At UCL, the longest TtTS reached 29.67 seconds, while at LU, the maximum was 12.67 seconds.

In the Congestion scenario, using a cradled mobile phone significantly increased the time it took participants to reach the target speed compared to when they were not engaged in any activity (No NDRA) ($p < 0.001$). Drinking water allowed participants to reach the target speed slightly faster compared to the No NDRA condition ($p < 0.05$). Using a handheld mobile phone significantly increased TtTS compared to the No NDRA condition ($p < 0.001$). The NDRAs “Eating popcorn”, reading a “Magazine” and doing a “Wordsearch” showed no significant difference from the No NDRA condition ($p = 0.32$, $p = 0.47$, $p = 0.23$ respectively), suggesting that these activities did not alter TtTS.

3.6 Reaction and performance: Lane deviations

Key findings

- There is no significant effect of specific NDRAs or location (LU vs. UCL) on lane deviation.
- Although UCL appears to have higher raw lane deviation percentages in many cases, these differences are not statistically significant, indicating that other factors might be influencing these outcomes.

Lane deviations in the form of swerving can indicate that the participant is struggling to control the vehicle and respond to unexpected environmental changes in the road which could occur if they lack SA or if they are being distracted still by an NDRA that has not been fully disengaged from.

A Generalized Linear Mixed Model (GLMM) was used to analyse the likelihood of lane deviation based on activity and scenario type, with location and participant treated as random effects to account for variability between locations and individuals. A binomial family was used for the model due to the binary nature of the lane deviation outcome (deviated or didn’t). Full model details and results are available in Technical Appendix 6.9.

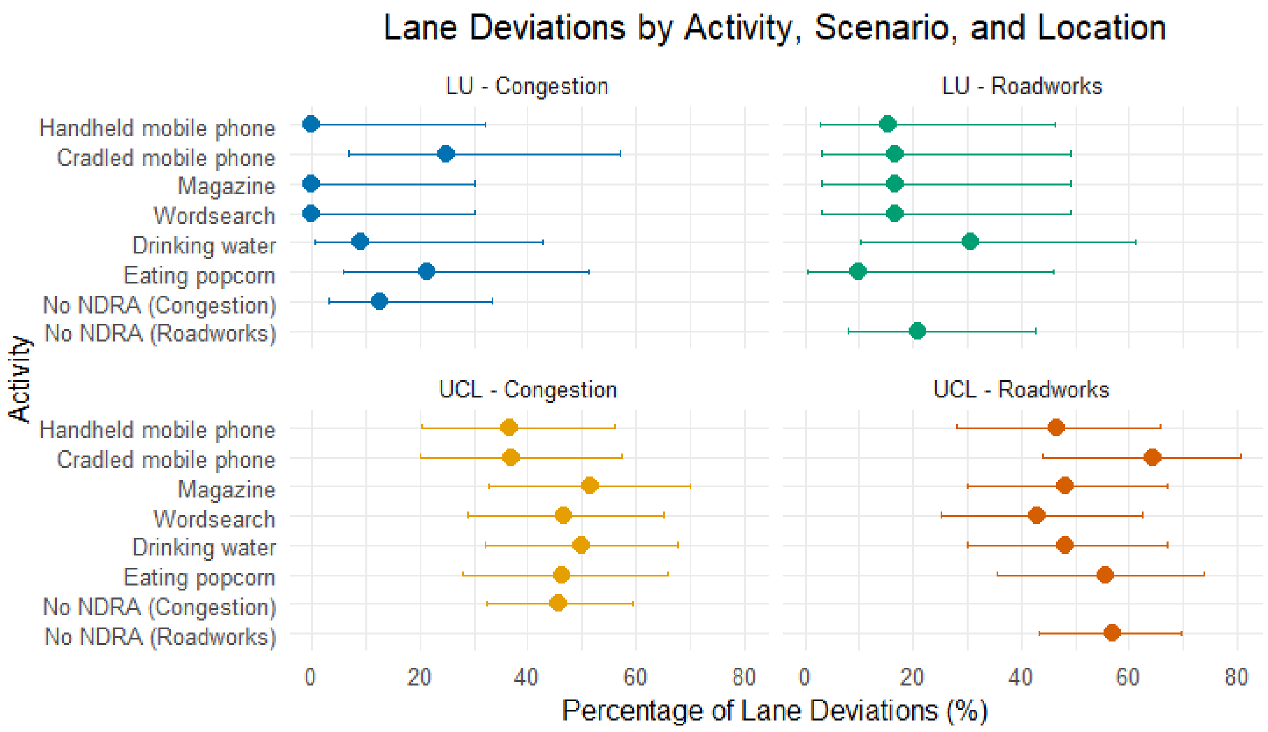


Figure 12: Percentage of lane deviations (swerving incidents) across NDRAs, scenarios, and locations.

Table 7: This table shows a comparison between location (LU and UCL) of the scenario Congestion in lane deviation for each NDRA

	LU		UCL	
Activity	**No Lane deviations (%)**	**Lane deviation (%)**	**No Lane deviations**	**Lane deviation (%)**
Cradled mobile phone	75.0	25.0	63.0	37.0
Drinking water	90.9	9.1	50.0	50.0
Eating popcorn	78.6	21.4	53.6	46.4
Handheld mobile phone	100.0	0.0	63.3	36.7
Magazine	100.0	0.0	48.3	51.7
No NDRA (Congestion)	87.5	12.5	54.4	45.6
Wordsearch	100.0	0.0	53.3	46.7

Table 8: Comparison between location (LU and UCL) for the scenario Roadworks in Lane deviation in each NDRA

	LU		UCL	
Activity	**No Lane Deviation (%)**	**Lane Deviation (%)**	**No Lane Deviation (%)**	**Lane Deviation (%)**
Cradled mobile phone	83.3	16.7	35.7	64.3
Drinking water	69.2	30.8	51.7	48.3
Eating popcorn	90.0	10.0	44.4	55.6
Handheld mobile phone	84.6	15.4	53.6	46.4
Magazine	83.3	16.7	51.7	48.3
No NDRA (Roadworks)	79.2	20.8	43.1	56.9
Wordsearch	83.3	16.7	57.1	42.9

There was no significant impact of any NDRAs on lane deviation. Figure 12 (and Tables 7 and 8) shows that in the "UCL - Roadworks" scenario, participants had higher proportions of lane deviations across most activities, particularly compared to the "LU - Congestion" scenario, where deviations were generally lower, but this difference is not statistically significant. For instance, in the "Cradled mobile phone" activity during the Roadworks scenario, lane deviation was observed in 64.3% of UCL participants compared to 16.7% of LU participants. However, this difference is not significant, suggesting that other factors, such as individual participant characteristics, might explain these variations rather than the location itself. The only factor approaching significance in the model is the scenario type, with Roadworks showing a trend towards higher lane deviations compared to Congestion ($p = 0.06$). This suggests that the Roadworks scenario might inherently be more challenging, leading to a higher likelihood of lane deviation, though the evidence is not conclusive.

3.7 Observation: Areas of Interest following a takeover request

Key findings

- Participants mainly focused on the "Road" and "Speedometer" after the takeover request.
- Right-side mirrors were checked more often than rear-view mirrors, but overall, mirror use was minimal following the takeover request and manual driving.
- Mirror checks increased during Congestion scenario and activities like reading a magazine or drinking water.
- Mirror use patterns were similar across both locations, showing a general tendency to neglect mirror checks.

The study compared the duration of looking times across various activities and scenarios, focusing on the period from the takeover request through the subsequent 30 seconds of manual driving. The initial analysis presents the mean looking times across all Areas of Interest (AOIs), to provide a comprehensive overview of participants' visual behaviour. A secondary analysis was then conducted, focusing on participants' use of only the mirrors across the different NDRAs.

A linear mixed-effects model (LMM) was fitted to assess the looking time across all areas of interest (AOIs) based on activity and area. The model treated activity and area as fixed effects, while location and participant were treated as random effects to account for variability across participants and locations. This comprehensive model helps to explore how different activities and areas impact looking time. Detailed results for this analysis are provided in Technical Appendix 6.10.

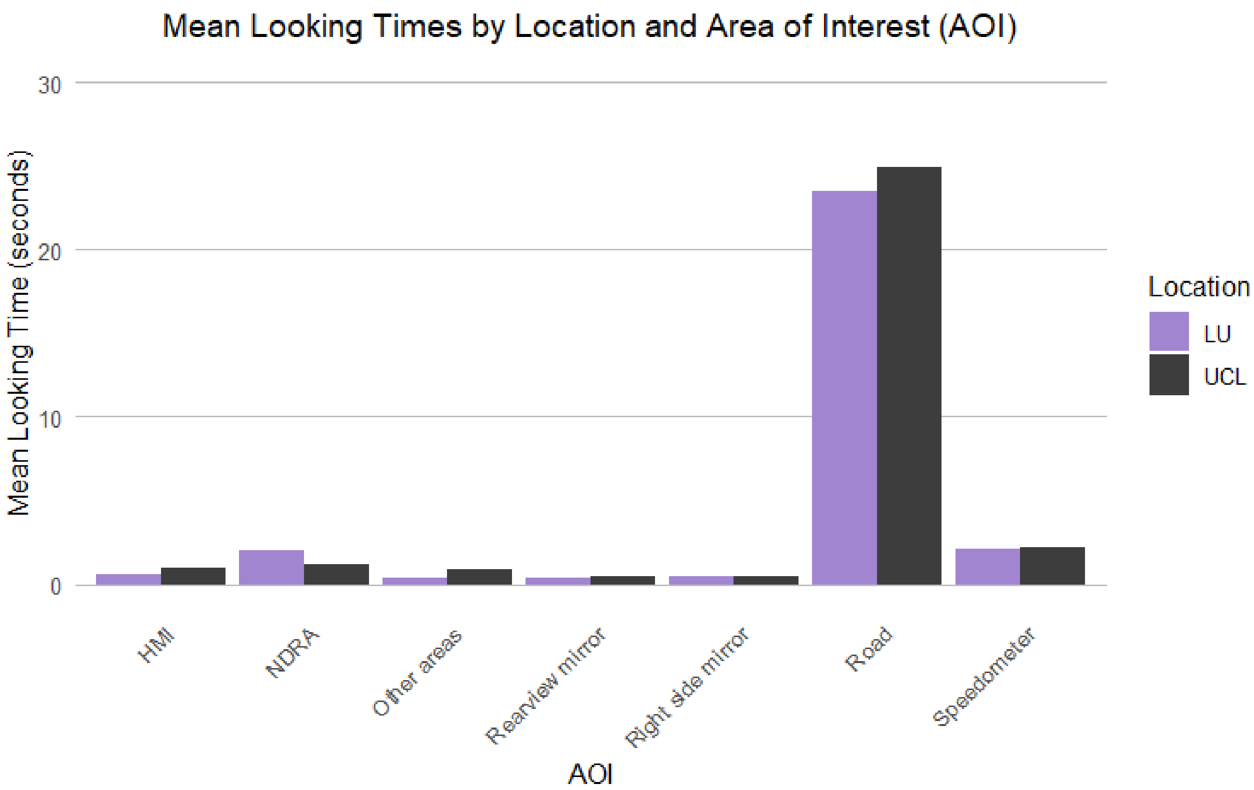


Figure 13: The bar chart illustrates the mean looking times across various areas of interest (AOI) by location (LU and UCL).

Table 9: This table displays the means and standard deviations of the duration of looking time (in seconds) across various areas of interest (AOI) for two locations, UCL and LU. The table is divided into columns for each

location, showing both the mean and standard deviation (SD) for each AOI. The AOIs listed include "HMI," "NDDT," "Other areas," "Rearview mirror," "Right side mirror," "Road," and "Speedometer."

Duration of Looking time (in seconds)				
	LU		UCL	
AOI	**Mean duration (in seconds)**	**Standard deviation (in seconds)**	**Mean duration (in seconds)**	**Standard deviation (in seconds)**
HMI	0.38	0.49	0.35	0.72
NDRA	0.91	1.44	0.53	1.25
Other areas	0.22	0.47	0.49	1.15
Rearview mirror	0.25	0.37	0.25	0.56
Right side mirror	0.31	0.49	0.36	0.63
Road	21.57	3.64	23.77	5.75
Speedometer	1.63	1.44	1.75	1.65

The analysis of looking times across different AOIs and locations (LU and UCL) reveals significant differences in visual attention distribution. The "Road" AOI dominated participants' attention, with mean looking times of 21.57 seconds at LU and 23.77 seconds at UCL ($p < 0.001$) (see Figure 13 and Table 9). The "Speedometer" AOI also received considerable attention, with mean looking times of 1.63 seconds at LU and 1.75 seconds at UCL ($p < 0.001$). This suggests that participants were mostly looking at the road ahead or monitoring the vehicle's speed.

The AOIs that participant looked at for the next longest durations were the HMI and the NDRA. The mean looking time for "HMI" was 0.38 seconds at LU and 0.35 seconds at UCL, showing that participants were searching for information, even though the HMI did not show anything other than the "Take control of vehicle". None of the other AOIs showed significant effects. The "NDRA" (Non-Driving Related Activity) AOI exhibited more variation between locations, with a mean looking time of 0.91 seconds at LU and 0.53 seconds at UCL ($p = 0.02$) which could mean that they were still engaging with the NDRA or looking back at it even while they were driving. The shortest durations of looking time were in the rear-view mirror in both locations. The next section focuses on looking times specifically in mirrors only.

Mirror looking time

A linear mixed-effects model (LMM) was used to examine whether mirror looking time differed by NDRA and AOI. This approach allows for an analysis of how different activities influenced the time participants spent looking in mirrors across various AOIs. Full details and results for this model are available in Technical Appendix 6.11.

Figure 14: The stacked bar chart displays the mean looking times for mirrors by activity and location (LU and UCL).

Table 10: The table presents the mean duration and standard deviation (SD) of looking times (in seconds) for the "Rearview mirror" and "Right side mirror" across various activities at two locations, LU and UCL. The activities listed include "Cradled mobile phone," "Drinking water," "Eating popcorn," "Handheld mobile phone," "Magazine," "No NDRA (Congestion)," "No NDRA (Roadworks)," and "Wordsearch." For each activity, the table provides the mean duration and SD for both LU and UCL, showing how long participants spent looking at the mirrors and the variability in their looking behaviour across the two locations.

Area	Activity	LU		UCL	
		Mean duration (in seconds)	**Standard deviation (in seconds)**	**Mean duration (in seconds)**	**Standard deviation (in seconds)**
Rearview mirror	Cradled mobile phone	0.15	0.24	0.21	0.37
	Drinking water	0.31	0.44	0.44	1.22
	Eating popcorn	0.25	0.46	0.17	0.38
	Handheld mobile phone	0.27	0.40	0.23	0.36
	Magazine	0.33	0.43	0.22	0.32
	No NDRA (Congestion)	0.33	0.39	0.27	0.46
	No NDRA (Roadworks)	0.20	0.29	0.19	0.39
	Wordsearch	0.18	0.28	0.25	0.40
Right side mirror	Cradled mobile phone	0.22	0.38	0.27	0.46

	Drinking water	0.33	0.54	0.30	0.48
	Eating popcorn	0.22	0.39	0.33	0.54
	Handheld mobile phone	0.28	0.53	0.35	0.52
	Magazine	0.42	0.50	0.48	0.74
	No NDRA (Congestion)	0.44	0.65	0.50	0.93
	No NDRA (Roadworks)	0.28	0.44	0.32	0.63
	Wordsearch	0.26	0.41	0.35	0.56

Participants generally spent more time looking at the right-side mirror compared to the rear-view mirror, regardless of the activity (see Figure 14 and Table 10), with this difference being statistically significant ($p < 0.001$). These differences may be attributed to the fixed placement of the right-side mirror at UCL on the vehicle itself, as opposed to its placement on the projector at LU. The mirror's positioning at UCL might have made it more convenient and intuitive for participants to check it frequently, leading to the observed increase in looking times. However, these differences were not statistically significant when factoring in the location variable ($p = 0.74$).

Participants looked in the mirrors significantly more during the "No NDRA (Congestion)" scenario compared to other activities ($p = 0.001$). Participants also looked in the mirrors more during the "No NDRA (Congestion)" scenario than during the "No NDRA (Roadworks)" scenario, with the latter showing no significant increase in mirror use ($p = 0.50$). This could suggest that participants perceived congestion as a more dynamic and potentially hazardous situation requiring more frequent mirror checks, whereas roadworks might have been seen as a more static or predictable environment, leading to less frequent mirror use.

Furthermore, the mean looking times for the right-side mirror were consistently higher at UCL across most activities compared to LU. For example, during the "Cradled mobile phone" activity, participants at UCL spent an average of 0.27 seconds looking at the right-side mirror, compared to 0.22 seconds at LU ($p<0.001$). Similarly, during the "Magazine" activity, the mean duration was 0.48 seconds at UCL, compared to 0.42 seconds at LU ($p<0.001$). Another significant increase in mirror use was observed when participants were drinking water ($p = 0.009$). This activity likely required participants to divert their attention from the road briefly to put the cup back in the holder, prompting them to compensate by checking mirrors more frequently once they resumed control.

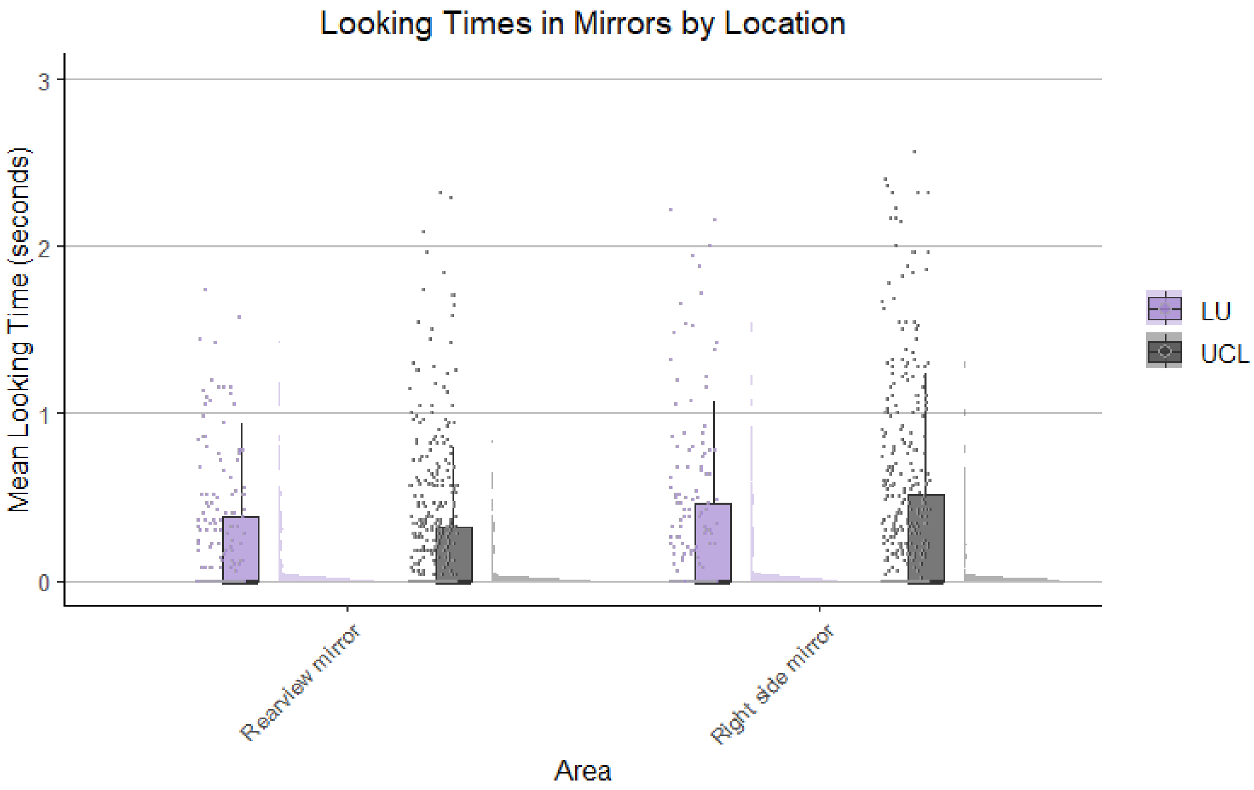


Figure 15: The raincloud plot shows the mean looking times in mirrors by location (LU and UCL).

The raincloud plot in Figure 15 illustrates that participants generally did not spend much time looking in mirrors after the takeover request, indicating a widespread tendency to neglect mirror checks. Most participants spent minimal time on this task, with a noticeable portion spending less than a second or not looking at their mirrors at all in any trials (1 participant at LU and 4 participants at UCL). The patterns of mirror use are similar across both UCL and LU locations, suggesting that this behaviour is consistent regardless of the location.

3.8 Observation: Pupil diameter change rate

Key findings

- Activities like using a handheld mobile phone, doing a wordsearch, or eating popcorn were linked to smaller changes in pupil size. This suggests that doing one of these NDRA could affect being ready to drive when the car switches to manual control.
- Drinking water also had a similar effect, though it was less pronounced. It still slightly reduced focus or alertness.
- Driving in congested traffic (No NDRA - Congestion) had the opposite effect—it caused bigger changes in pupil size. This suggests that when drivers were taking over in congestion, they were more alert and focused, likely because the situation required more attention.

As previously explained in Section 3.2, the rate of change in pupil diameter serves as an indicator of SA and cognitive load during the takeover process. This measurement helps gauge how drivers react to regaining control from an automated system. A higher rate of pupil diameter change typically reflects increased cognitive load or stress, indicating that the driver is actively working to regain full awareness of the driving environment. In contrast, a lower rate of change might suggest that the driver is more relaxed or less engaged.

We fitted two linear mixed-effects models to analyse pupil diameter change rate. The first model assessed the effect of scenario type and NDRA while accounting for variability in location and participants (Appendix 6.12). The second model introduced an interaction between NDRA and location to determine if the relationship between NDRA and pupil diameter change rate differed by location (Appendix 6.13). This explored whether the impact of activity on pupil responses varied across different locations.

Activities like "No Task (Congestion)" show high cognitive load at UCL but significantly lower at LU, and NDRAs such as reading a "Magazine" and "Wordsearch" exhibit higher cognitive load at LU compared to UCL. Significant variability in cognitive load across activities between UCL and LU complicates drawing definitive conclusions (see Figure 16).

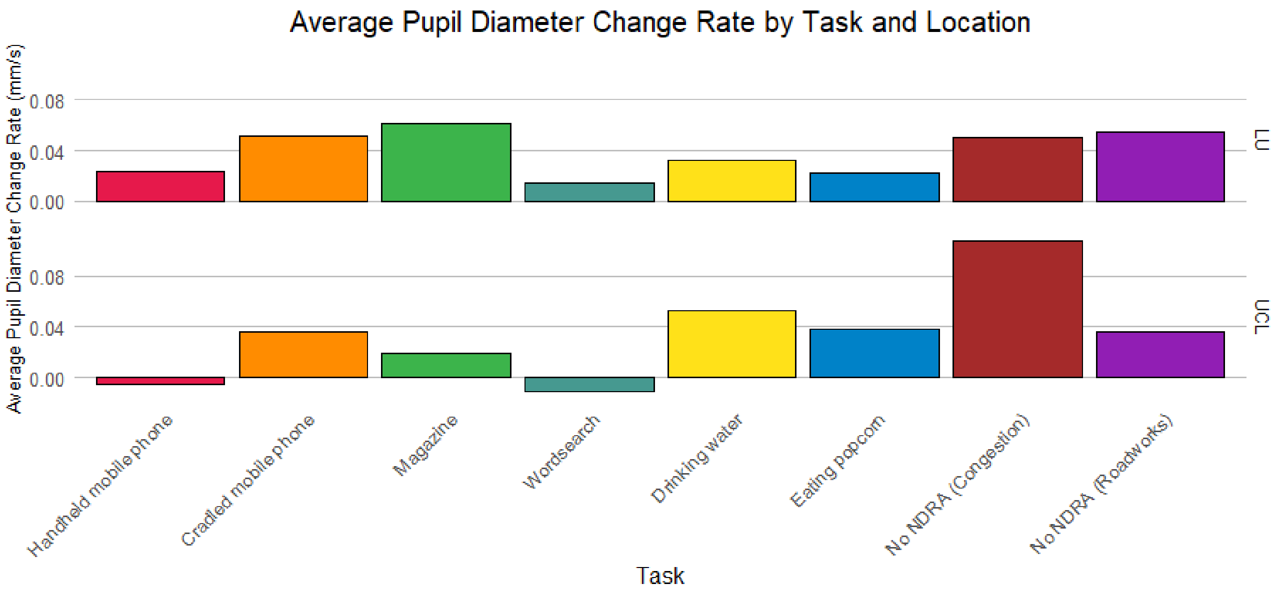


Figure 16: This interaction plot illustrates how different activities affect the pupil diameter change rate, which is a measure of cognitive load, across two locations: UCL and LU.

Certain activities, such as using a handheld mobile phone, doing a wordsearch, and eating popcorn, significantly decreased the pupil diameter change rate ($p < 0.001$ for each NDRA respectively). This suggests that these tasks might lower cognitive load or engagement when taking over control from a self-driving car, potentially reducing SA. Drinking water also caused a slight but significant reduction in pupil diameter change rate ($p < 0.05$), indicating that even motoric tasks can subtly influence a driver's readiness to regain full SA.

In contrast, engaging in the No NDRA (Congestion) activity led to an increase in the pupil diameter change rate ($p < 0.001$), reflecting a higher cognitive load or stress during the transition from automated to manual driving in congestion scenarios. This suggests that drivers are more alert or stressed in such situations, which may reflect heightened SA needed to manage the complex driving conditions. However, activities like reading a magazine and driving in No NDRA (Roadworks) did not significantly affect the pupil diameter change rate, indicating that these activities may not substantially impact SA during the takeover process.

3.9 Reflection: Subjective Measures of Situational Awareness

Key findings

- No significant difference in self-reported situational awareness was found between activities for any participants.
- Perceived workload, as measured by NASA-TLX, showed no significant differences across activities, despite some informal feedback suggesting that eating popcorn felt more difficult.
- Small differences in situational awareness were observed between the Congestion and Roadworks scenarios, but no significant differences were found between tasks.

The NASA-TLX workload index and the SART questionnaire were used to evaluate perceived workload and subjective measures of situational awareness.

NASA TLX

Participant responses for NASA-TLX showed few variations across NDRAs. Although participants made informal comments about high workload for performance of the Eating Popcorn activity, (as the design of the activity made it a more difficult action than normal eating) there were no significant differences found in workload between NDRAs for any NASA dimensions.

SART: Measures of Situational Awareness

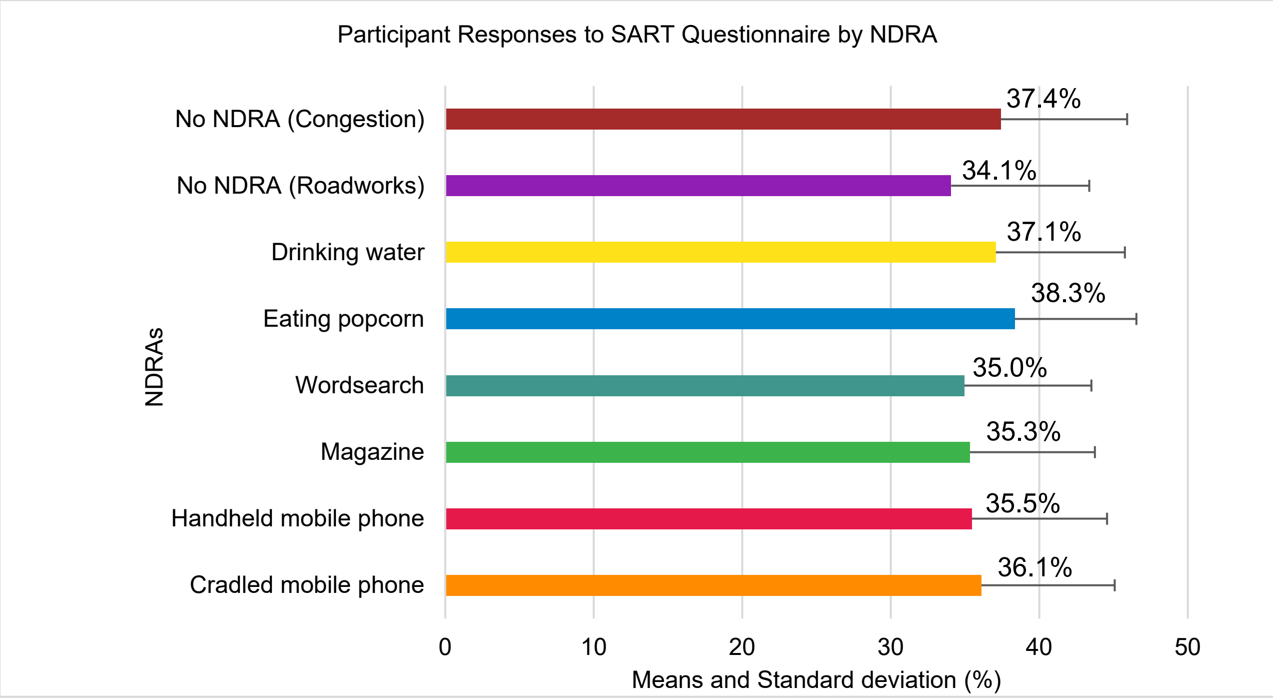


Figure 17: "Mean and Standard Deviation of Participant Responses to SART Questionnaire by Activity," which displays the mean and standard deviation of participant responses (expressed as a percentage) to the SART questionnaire across various activities.

SART scores suggest that simpler activities like eating popcorn and drinking water allowed participants to subjectively feel that they had higher situational awareness after the takeover request for these NDRAs (see Figure 17). In contrast, activities that require more cognitive and visual engagement, such as mobile phone activities and non-technological activities had lower subjective scoring of SA. The small difference in SART scores between the No Task Congestion and Roadworks scenarios highlight the inherent complexity and variability of the Roadworks scenario, which naturally reduces SA even without additional activities. Statistical tests showed no significant differences in self-reported SA for any NDRAs for all participants.

3.10 Survey data

Pre-experiment questionnaire

We asked participants two key questions in the pre-experiment questionnaire: one regarding recent purchases/subscriptions to tech products and services, and the other about their comfort with various transport innovations to ascertain their openness to technology (see Figure 18).

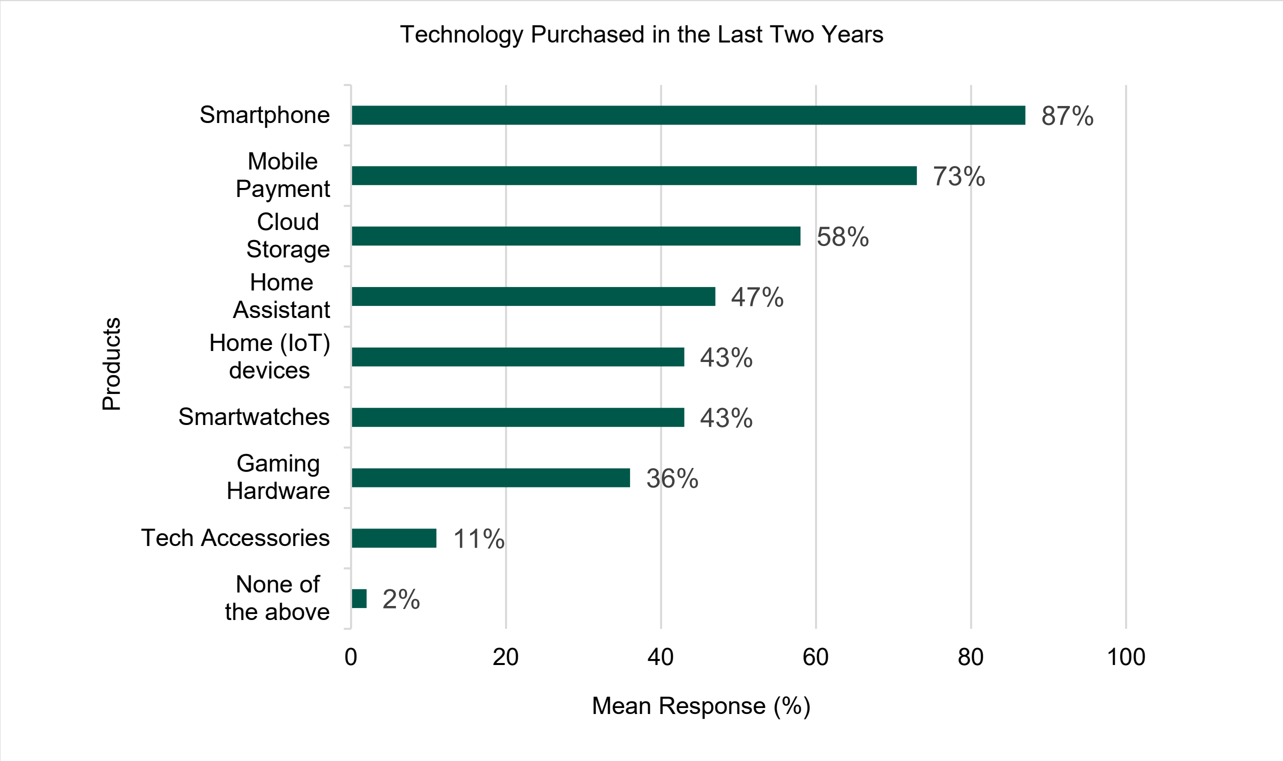


Figure 18: We asked, '*Did you purchase any of the following products / subscribe to any of the following services over the last two years?*'.

A significant majority of participants (87%) had purchased a smartphone within the last two years, reflecting a strong openness to new technology. High adoption rates were also seen with mobile payments (73%) and cloud storage (58%), indicating that this group is comfortable with digital solutions and likely to be ready for advanced vehicle technologies. Additionally, 47% of participants owned home assistants, and 43% owned smartwatches or IoT devices, showing a substantial interest in smart and connected living, which may enhance their adaptability to self-driving car features. Only 2% did not purchase or subscribe to any of the listed items, which might indicate either infrequent updates to their digital assets or a preference to avoid technology. Overall, the sample demonstrates a broad integration of technology into participants' lives, suggesting a smooth potential transition to self-driving car systems.

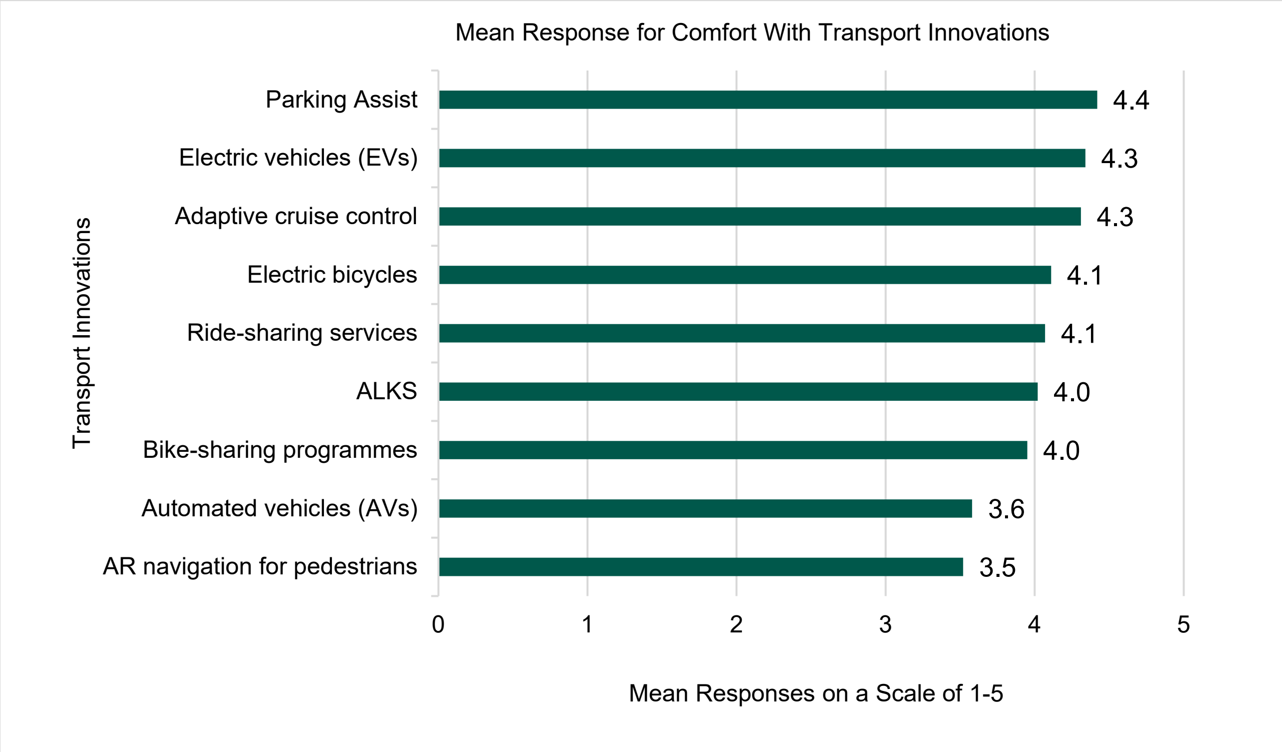


Figure 19: We asked, '*How comfortable are you with accepting the following transport innovations in the future?*', ranked by a scale of 1 to 5, where 1 indicates 'Not comfortable at all' and 5 indicates 'Very comfortable'.

Established and practical transport innovations were widely accepted (shown in Figure 19), while the trend tipped towards cautious optimism for fully automated and highly novel technologies. Participants showed high comfort levels with parking assist (4.4) and Electric Vehicles (4.3), technologies that are already widely available and well-understood. Adaptive cruise control (4.3) and Automated Lane Keeping Systems in vehicles (4.0) are viewed favourably, suggesting trust in incremental advancements that enhance driver assistance without fully removing control. High comfort with e-bikes (4.1), ride-sharing services (4.1), and bike-sharing programmes (4.0) reflects an openness to sustainable and shared transport modes, likely influenced by their tech-savvy and eco-conscious mindset. However, lower comfort levels with fully automated vehicles (3.6) and Augmented Reality navigation for pedestrians (3.5) indicate a need for further development, demonstration of reliability, and building user trust in these emerging technologies.

Post- experiment questionnaire

We asked participants to rate their level of immersion in the NDRAs while the self-driving was engaged in the study. The trends in Figure 20 indicate that activities involving high cognitive and visual engagement, such as mobile phone activities and non-technological activities such as the Wordsearch, were reported as the most immersive. Participants often became deeply engrossed in these activities, potentially reducing their SA when they come to takeover. Conversely, simpler activities like eating and drinking were less immersive, allowing participants to remain more aware of their surroundings.

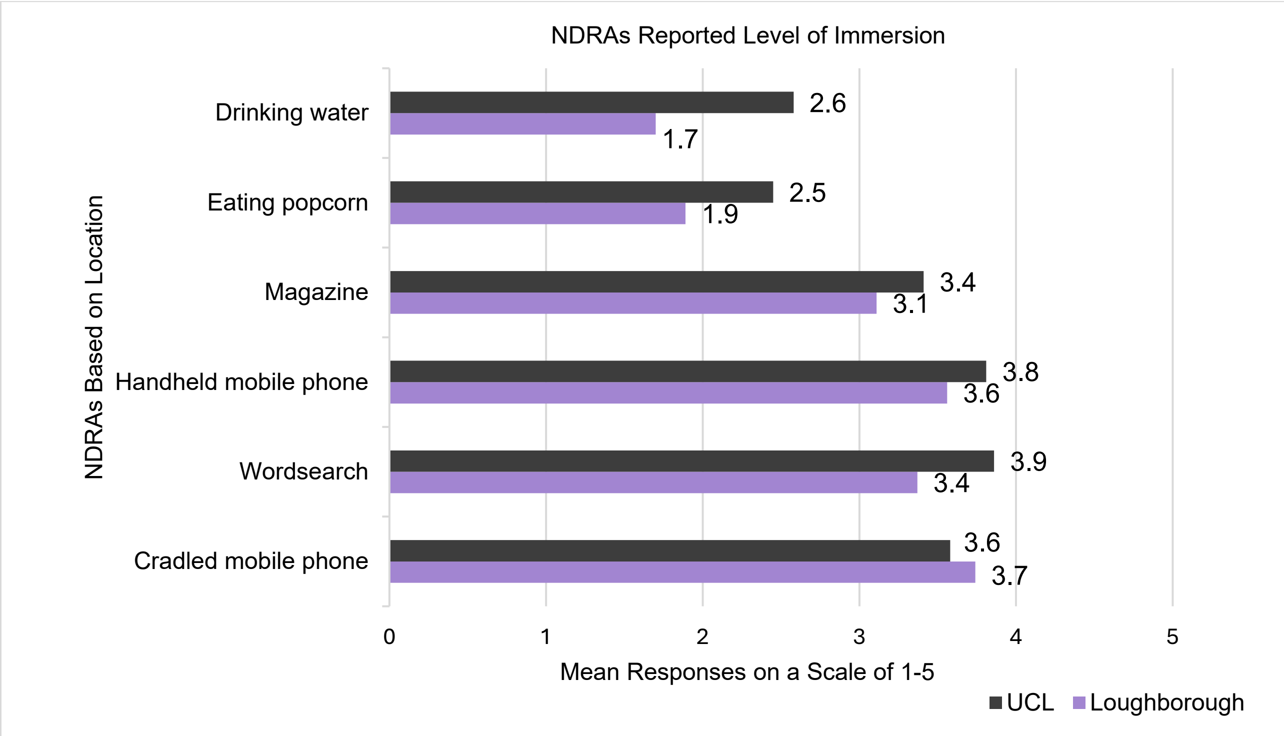


Figure 20: We asked, 'Which non-driving related activities did you find more or less immersive while operating the CAV system?'. Below is the average level of immersion, shown in rank order, where 1 indicates 'Not immersed at all' and 5 indicates 'Completely immersed'.

Participants were also asked their opinion on whether any NDRAs should not be allowed while using a self-driving vehicle as a UiC (see Figure 21). The responses suggested that there was significant concern about activities that require high cognitive engagement, like playing handheld games and completing puzzles. Many participants felt these activities should be prohibited due to their potential to distract drivers from the primary activity of monitoring the driving environment. On the other hand, low-engagement activities like eating and drinking were generally not considered problematic. Some participants (4% at LU and 20% at UCL) after their experience felt that no activities should be allowed at all!

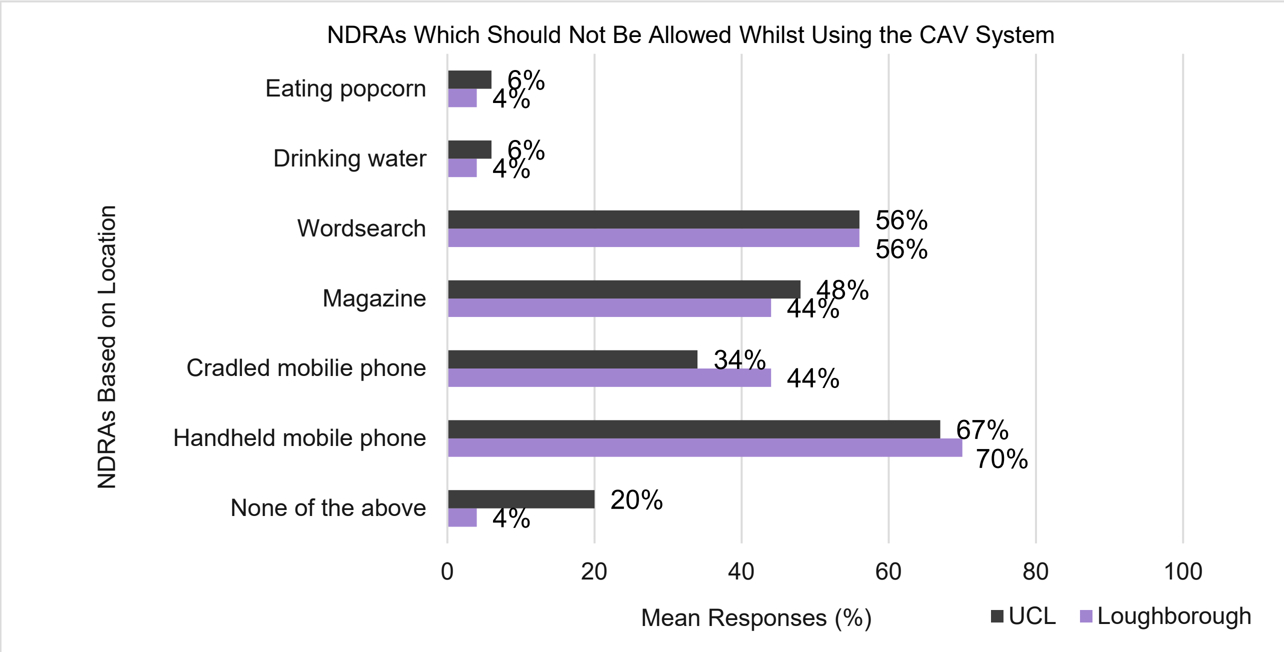


Figure 21: We asked, 'Do you think any of the NDRAs should *not be allowed* while using the CAV system?'

Finally, when asked about enjoyable or challenging aspects of the experiment setup and procedures, participants highlighted that high-engagement activities often made it difficult to maintain SA, suggesting that they felt they were supposed to be doing this even when the self-driving was engaged. Many reported feeling less 'aware' while engaged in immersive activities, indicating that these activities could significantly detract from their ability to takeover effectively.

Participants also shared insights into specific moments during the experiment when they felt more or less aware of their surroundings.

"After the phone and Tetris activity, you kind of forget where you are."

"I felt less aware when doing activities that required mental and physical effort combined."

"I was more aware on the simpler activities. Heavily aware on alert beeping".

"I felt more aware when not doing an activity".

"As I was aware of the test situation, I maintained an awareness for the most part however it is quite easy to lose concentration with time."

These comments suggest that participants could be confused as to what they were supposed to be doing during the periods of self-driving and that they should have been, in some part, situationally aware. This could be a result of the lack of direct instructions at the start of the study, which were deliberately vague to enable an exploration of natural behaviour.

4. Discussion

4.1 Key findings

This study investigated the impact of various non-driving related activities (NDRAs) on a User-in-Charge's (UiC) situational awareness (SA) during the transition from automated to manual control in self-driving vehicles after a planned takeover request.

Some NDRAs were easier to disengage from than others impacting takeover time. A critical factor influencing the effectiveness of the takeover was how quickly participants disengaged from the NDRA. For instance, activities like reading a magazine resulted in quicker and more consistent takeover times, likely because participants could more easily disengage from them. In contrast, tasks involving mobile phones were associated with low disengagement rates, where participants often initiated manual driving without fully stopping the activity. This led to shorter takeover times but did not necessarily constitute a safe transition, as participants might not have fully understood what dynamic driving actions were required for the driving task relevant to the scenario.

Ability to safely takeover within regulated timeframes was variable. Participants were not told about the ALKS regulations of taking over within 10 seconds but still seemed to be trying to take over as quickly as possible. These findings suggest that while some participants managed to take control within the required timeframe, they might not have done so in a manner that ensures safe driving. Many other participants across both sites took far longer than 10 seconds; this could be because they were carefully disengaging (some people even put the lid back on the pen) and looking around to build SA or they were just naturally slower. Either way, this variability raises concerns about the adequacy of the 10-second takeover period mandated by ALKS regulations (United Nations, 2021). Even seemingly simple tasks can negatively affect a driver's readiness to assume control, highlighting the broad impact of NDRAs on takeover performance.

Driving scenario complexity impacts takeover times. The context of the driving scenario also played a role in takeover times, with roadworks generally resulting in slower and more variable responses compared to congestion. This suggests that participants might have struggled to pick up on critical cues from the environment, such as road signs or changes in traffic flow, which are essential for adapting their driving speed and maintaining control. The increased difficulty in the roadworks scenario implies that different NDRAs might have varying impacts depending on the context. For instance, tasks like using a cradled mobile phone, eating popcorn, drinking water, and doing a wordsearch significantly delayed the time to reach target speed in the roadworks scenario, while these same activities did not have as pronounced an effect in the congestion scenario. However, the observed differences between the two scenarios were minimal, indicating that while the takeover scenario does influence how quickly participants can regain control, the actual difference in time is very small—less than half a second. Even this seemingly minor delay can be dangerous; at motorway speeds of 70 mph (about 113 km/h), a vehicle can travel approximately 15.4 metres (around 50 feet) in just half a second. This distance could be critical in avoiding potential hazards.

Individual differences and environmental factors play a significant role in takeover performance and SA. High levels of variability were evident between participants and across locations (UCL and LU), suggesting that individual differences and environmental factors play a significant role in takeover performance and SA. UCL participants generally took longer and showed greater variability in reaching target speed, particularly in the roadworks scenario. Activities such as using a cradled mobile phone, eating popcorn, drinking water, and completing a wordsearch significantly delayed TtTS at UCL compared to LU, highlighting the potential for certain NDRAs to impair the driver's ability to adjust to the context of the driving task quickly and accurately. Analysis of lane deviations as a metric of driving performance revealed no significant effects of specific NDRAs. However, again, observed variability across different activities and locations underscores the importance of considering individual and contextual factors in assessing takeover performance.

Mirror checks were rarely used to build SA following a takeover request. Eye-tracking data revealed that while participants primarily focused on the road and speedometer, mirror checks were minimal, indicating a potential gap in fully assessing the driving environment. Additionally, there were noticeable searches directed towards the HMI, indicating that participants may have been looking for information to build SA but were uncertain where to find it. This points to a potential area for improvement: the HMI could be better utilised to support effective takeovers by providing clearer information about the reason for the takeover request. By enhancing the HMI to deliver more context-specific details, participants could build SA more quickly, leading to safer and more effective transitions from automated to manual driving. Furthermore, changes in pupil diameter suggested varying levels of cognitive load depending on the activity, with more alert responses observed during congestion scenarios.

Clearer guidance on what constitutes a safe and effective takeover is needed. Although there is no research on the optimal way to build appropriate SA prior to taking manual control, it stands to reason that guidance would include taking the necessary time to safely store the NDRA and perform essential actions such as mirror checks before resuming control. However, many participants either struggled to properly disengage from the task or did not recognise that continuing to hold onto the NDRA or allowing it to continue (e.g., not pausing a mobile phone activity) compromised the safety of the takeover. This observation is not entirely surprising, given that instructions to participants were deliberately vague: "take over as soon as you feel it is safe to do so." This ambiguity likely contributed to the varied interpretations of what constitutes a "safe" takeover. While this approach allowed the observation of naturalistic behaviours with the interactions of the NDRA, the results strongly suggest that a UiC of self-driving vehicles will require clearer guidance on what constitutes a safe and effective takeover. Specifically, they need explicit instructions on their responsibilities during the transition, such as the importance of disengaging fully from any ongoing activities and how best to build the necessary SA before taking back manual control.

4.2 Comparison with previous research

Impact of NDRAs on driving performance

In this study, the longest takeover times were associated with tasks such as using a handheld mobile phone or engaging in a word search, consistent with previous research by Eriksson and Stanton (2017) and Mok et al. (2017), where takeover times typically ranged from 3 to 8 seconds. However, raw data revealed even greater variability, with the longest recorded takeover time in this study reaching up to 24.4 seconds for reading a magazine. This is longer than the upper range reported in earlier studies, suggesting that the specific nature of the NDRA, as well as contextual and environmental factors, can substantially extend the time required to regain control of the vehicle. Interestingly, Merat et al. (2014) found that drivers required approximately 40 seconds to fully stabilise and regain control after automation, which suggests that even if a takeover occurs, drivers may still be in the process of building SA and may not yet be fully prepared to drive safely. Although most participants in this study were able to take over within the ALKS 10-second period, it does not necessarily mean they were doing so safely if they had not fully disengaged from the NDRA and were still distracted.

The analysis of lane deviations revealed no significant impact of NDRAs or location on lane-keeping performance, which diverges somewhat from studies like those by Zeeb et al. (2016), who found that certain NDRAs could impair lane-keeping performance, indirectly indicating lower SA. Although there was a trend towards higher lane deviations in the roadworks scenario at UCL, these differences were not statistically significant. This suggests that while scenario complexity might influence driving performance, it may not be as strongly linked to SA as previously thought, or that other unaccounted factors could be influencing these results. The variability between the two testing locations (UCL and LU) further highlights the potential impact of situational factors on driver performance, indicating that the context in which the takeover occurs may influence the time needed for drivers to resume control more than previously acknowledged. This variability could point to the influence of situational factors not fully explored in earlier research.

NDRA Engagement and Its Effect on Situational Awareness

Shaw et al. (2020) observed that many drivers, after receiving a takeover request, briefly glanced at the road but then reverted to their previous activity, delaying their readiness to assume full control of the vehicle. There was a similar pattern in this study, particularly with mobile phone use, where participants often started manual driving without fully disengaging from the task. However, although participants at LU displayed more diligent behaviour in putting the activity away, it led to significantly longer takeover times, sometimes exceeding 10 seconds. This suggests that if a UiC takes the time to properly disengage, it may negatively impact takeover times.

Large, Burnett, and Salanitri (2019) noted that participants in driving simulator studies often struggled to shift their focus back to the driving task upon receiving a takeover request. Findings from this study concur, particularly in tasks that required high levels of engagement.

If participants had been allowed to choose their own tasks, as in their study, they might have found it even more challenging to disengage and start building SA. This highlights the importance of providing clear instructions to participants, a factor we could improve in future studies.

Effect of types of NDRA on Situational Awareness

Du et al. (2020) found that hands-free activities, such as watching a film on a tablet, allowed drivers to quickly shift their attention back to the driving task, typically within 2 seconds. This contrasts with the study findings, where participants engaging in similar tasks (such as watching a film on a cradled mobile device) often experienced longer takeover times, especially when the task was more immersive or required manual interaction, like using a handheld mobile phone to play a game. Results from this study suggest that the physical involvement and cognitive load associated with handheld devices can delay the takeover process, although this effect was also observed with non-technological tasks like reading a magazine or completing a wordsearch.

Jiang, Wang, and Tang (2024) found that gamified attention activities during simulated driving scenarios could enhance reaction times and decision accuracy during takeovers. Contrary to these findings, this study showed that such tasks, including Tetris, resulted in slower takeover times and poorer disengagement. This discrepancy could be due to the handheld nature of the task, which may have hindered participants' ability to quickly switch focus. Exploring the impact of relocating such activities to the vehicle's HMI, where they could automatically freeze upon a takeover request, might yield different results.

Vogelpohl et al. (2018) suggested that while drivers can deactivate automation quickly, more time is needed to build SA and respond effectively to unexpected traffic events. Although this study did not directly investigate unplanned scenarios, the findings align with the notion that building SA after a takeover request requires additional time, particularly when participants are engaged in complex NDRAs. Lu et al. (2017) found that it takes between 7 and 20 seconds to develop sufficient Level 1 SA after engaging in complex tasks. While this study used different tasks, we observed a similar trend where more complex activities resulted in longer takeover times, suggesting that regaining SA is a time-consuming process. Radlmayr et al. (2014) highlighted the differential impact of NDRAs on SA, finding that some activities could impair the ability to regain SA after a takeover request. Results from this study agree, showing that more cognitively demanding tasks, like interacting with a mobile phone, completing a wordsearch or reading, were associated with longer takeover times than drinking water.

Observation and Situational Awareness

In terms of eye tracking and visual attention, the study findings align with those of Kunze et al. (2019) and Liang et al. (2021), who emphasised the importance of visual scanning in rebuilding SA after a takeover request. In this study, participants focused mainly on the road and speedometer after the takeover request, indicating basic Level 1 SA (perception of the road ahead). However, there was a notable neglect of mirror checks, which are crucial for develop Level 2 and Level 3 SA (comprehension and projection of future events). This finding suggests that while drivers may quickly refocus on essential driving responsibilities, they may not fully regain the broader awareness needed for safe driving.

Miller et al. (2015) found that engaging in NDRAs, such as watching videos or reading on a tablet, could help maintain arousal and prevent drowsiness, potentially benefiting takeover performance. However, findings in this study did not support this finding, as activities like using a handheld mobile phone, doing a wordsearch, or eating popcorn were linked to smaller changes in pupil diameter, suggesting that these activities might reduce readiness to take manual control. Conversely, we found that the driving scenario itself, particularly congested traffic, had a more significant impact on arousal, as indicated by larger changes in pupil size, which suggests heightened alertness and focus during these scenarios.

4.3 Limitations of the study

Findings, while providing valuable insights into the impact of NDRAs on SA and driver performance, should be interpreted within the context of the study's limitations.

Differences between sites

A significant limitation of this study was the procedural discrepancies between the two testing sites, LU and UCL. At LU, participants were required to use a verbal "Ready to drive" command to regain manual control, whereas UCL participants took control by directly engaging the steering wheel and pedals. This procedural difference likely introduced variability in takeover times, potentially influencing the results. The verbal command at LU added an extra step, which could have delayed the transition to manual control and impacted situational awareness and takeover performance. Additionally, the placement of mirrors differed between the sites, with LU using mirrors projected onto screens, while UCL had fixed mirrors on the vehicle itself. Despite this, the study found minimal differences in visual attention between the two locations, suggesting that the environmental setup may be less impactful than individual driver behaviour.

Some participants at UCL reported difficulty in controlling the driving simulator, noting that the steering wheel felt either too light or too stiff compared to their own vehicles. These initial challenges could have affected performance, particularly in the early stages of the experiment. However, most participants quickly adapted after the initial training period, and no significant issues were reported thereafter. Future studies should standardise takeover procedures and driving simulator setup across sites to ensure consistent data collection, with a recommendation to implement direct input controls at LU to minimise delays.

Small sample sizes

Smaller sample sizes, particularly at LU where data loss and non-compliance reduced the number of participants from 31 to 24, likely contributed to increased variability and sensitivity to outliers, making the findings less robust compared to the larger sample at UCL. The reduced sample size at LU may have led to a stretched distribution and greater inconsistency in results, as observed in the wider spread of the density curves. Differences in protocols between the two sites also limited the ability to combine data. Future research should aim for larger sample sizes at both locations to enhance the robustness and comparability of the findings.

Sample restrictions

Sample recruitment excluded participants who wore glasses. However, as individuals age, the likelihood of needing corrective eyewear for both reading and driving increases. Many people use different types of glasses for these activities, such as bifocals, progressive lenses, or separate pairs for near and distance vision. The process of switching between these glasses or adjusting to a single pair with multiple prescriptions could introduce a delay in the takeover process.

Additionally, the need to adjust focus when switching between reading glasses and distance vision could impact SA during a takeover. For example, a driver may need to transition from reading a screen or printed material (requiring near vision) to scanning the road and mirrors (requiring distance vision). This delay could be critical, as it might affect a driver’s ability to regain control of the vehicle quickly and effectively, especially in situations requiring rapid responses.

Given these factors, future research should include participants who wear glasses to better understand how these variables influence takeover performance. Studies should specifically explore the impact of different types of eyewear on the speed and quality of takeovers and whether certain types of glasses or corrective measures might mitigate any adverse effects. This would provide a more comprehensive understanding of the challenges faced by a significant portion of the driving population, particularly as automated driving technology becomes more widespread.

Lack of realism of some NDRAs

Including motoric NDRAs such as eating and drinking in the study was essential to examine how cognitively undemanding tasks that still occupy the hands could affect the takeover process. These activities may influence the transition back to manual control differently than more cognitively engaging tasks, as they require physical actions that could delay the rebuilding of SA. However, the realism of the eating task used in the study could be questioned. The task was designed to mimic the act of eating, but participants could have encountered difficulties as the simulated task involved fake food items and unfamiliar apparatus. Many people already eat and drink while driving, so it may have been more insightful to include a more complex and less typical NDRA, such as using a knife and fork or chopsticks, which is not feasible during manual driving. This would provide a more accurate assessment of how challenging motoric tasks impact SA and takeover performance in automated driving contexts. Alternatively, the tasks themselves may not have fully captured their attention. Drawing from Burnett et al. (2019), it might be more effective to allow participants to choose their own tasks to ensure sufficient immersion. Future research should consider incorporating such task changes to better understand their effects.

Eye tracking measurement

The eye-tracking data collected covers the moment of the takeover request and the subsequent 30-second manual drive, providing insights into how participants began to rebuild SA. However, it is important to acknowledge that SA might not have been fully regained until well after participants had resumed manual control, and this limitation is something we plan to address in future phases of research (as suggested by Marti, 2022). The data in this study shows that participants did not spend any significant portion of the 30-second post-TOR period looking at mirrors, which is a concerning finding. Future research should incorporate a more granular time series analysis, tracking the time to the first fixation on mirrors and correlating this with disengagement from the NDRA. This would provide a clearer picture of the sequence of visual attention shifts leading up to and following the takeover, helping to better understand the process of rebuilding SA and ensuring safer transitions from automated to manual driving.

Participant instructions

Many participants at both sites expressed uncertainty about the tasks, asking questions such as how many sips of water they should take or what they should do with the task when a takeover request was issued. As the aim was to observe natural behaviour, the response to these queries was consistently, “as soon as you feel ready and safe to do so." However, eye-tracking data suggests that participants may have been unsure of their responsibilities during the self-driving periods. Although participants were informed during the welcome briefing that they could engage in a task during these periods, it was not emphasised that they were legally not required to monitor the road. As a result, some participants might have felt obligated to keep an eye on the road, leading to less immersion in the task.

Providing participants with a more detailed briefing that outlines the legal responsibilities of the UiC could help ensure that their behaviour during the study more closely mirrors real-world interactions with self-driving vehicles. By clarifying these responsibilities, participants would have a better understanding of what is expected of them in actual self-driving scenarios, leading to more authentic and relevant behaviour during the experiment.

Scenario design

A critical aspect of the study was to determine whether participants used environmental cues, such as road signs, to build SA before resuming manual control. This analysis primarily focused on whether participants looked at the signs during the roadworks scenarios. However, given that the same scenarios were repeated with different takeover times, participants might have already processed the information from previous encounters with the signs and thus did not feel the need to look at them again. Alternatively, participants may have gathered sufficient information from the initial signs, making it unnecessary to check subsequent ones, such as the third overhead gantry.

The fact that many participants slowed down suggests that they may have been situationally aware, even if they did not repeatedly check the signs. Future research should consider whether participants can see and process the signs effectively on the first encounter, and whether this negates the need to check them again in subsequent scenarios. To address these questions, future studies could redesign scenarios so that the content of the sign changes (40mph, 50mph etc), allowing for a more accurate assessment of whether participants' responses vary based on the new information. Including unexpected scenarios, such as sudden weather changes, can prevent participants from learning and anticipating the scenarios, thereby providing a more realistic assessment of their responses and SA.

Lack of baseline information

No NDRA trials were used as a baseline for assessing how participants took over control when they had not been engaged in any activity during the automated driving phase. However, without knowing participants' normal driving behaviour, it's difficult to determine whether their performance issues were due to a lack of SA or simply poor driving skills. A more effective approach would be to include a baseline drive where participants fully control the vehicle themselves before any automated driving or NDRAs are introduced. This baseline drive would allow researchers to measure participants' typical driving performance, including their standard response times, SA, and vehicle control. Establishing such a baseline would offer a crucial point of comparison, making it easier to understand the impact of automated driving and NDRAs on driving performance.

Length of self-driving periods

The study design included multiple trials with relatively short takeover requests, none of which lasted longer than 4 minutes. This brevity could have led participants to anticipate the takeover requests, potentially influencing their responses and not accurately reflecting real-world driving conditions, where drivers might experience more extended periods of automated driving between takeover requests. In practical driving scenarios, these periods would likely be longer, and the situations more complex, requiring drivers to maintain SA over extended durations. Future studies should consider increasing the duration of takeover requests to better simulate real-world conditions, providing a more accurate understanding of how drivers manage transitions from automated to manual control over longer periods.

Limitations of driving simulators

Certain aspects of both driving simulators lacked realism, which may have influenced participant behaviour and SA. For instance, the absence of real-world sensory feedback, such as vibrations from the road, could alter how participants perceive and react to the driving environment. Incorporating more realistic elements into the simulator, such as tactile feedback and enhanced environmental cues, could help mitigate these limitations. A general limitation of simulators is the reduced sense of "risk" compared to real-world driving, leading participants to potentially behave differently than they would in actual driving scenarios. However, it was observed that most participants seemed to attempt to drive as they normally would.

Although these limitations may affect the generalisability of the findings, the controlled environment provided an opportunity to isolate and examine specific variables related to non-driving-related activities (NDRAs) and takeover performance. While the results should be interpreted with some caution, particularly in their application to on-the-road driving, the insights gained offer valuable contributions to understanding driver behaviour in automated driving contexts. Further research conducted in real-world conditions is necessary to validate these findings and explore their broader implications.

4.4 Recommendations for future research

Accessibility and the role of the User-in-Charge

One area that requires further investigation is how the concept of a UiC interacts with individuals who have additional accessibility needs, such as those requiring glasses, experiencing hearing loss, or having broader disabilities. For example, as people age, the likelihood of needing glasses for driving or reading increases. Many individuals must switch between different types of glasses depending on the task—driving glasses for distance and reading glasses for close-up work. This switching process can introduce delays in the takeover process, which this study did not account for, as participants who wore glasses were excluded. Additionally, hearing impairments, whether age-related or otherwise, could impact the ability to hear the auditory takeover signal while engaged in a visual NDRA. It's important to understand how switching between visual aids or hearing impairments might delay the takeover process and affect the ability to regain SA and safely resume control of the vehicle. Future research should explore the time delays associated with visual or auditory accessibility differences and how these factors influence overall safety, particularly for older drivers who may face these challenges more frequently.

Integration of NDRAs within the Human-Machine Interface (HMI)

In this study, disengagement from NDRAs was entirely at the discretion of the participants, leading to varied and sometimes prolonged takeover times. This raises an important question: how might takeover performance be affected if NDRAs were integrated into the HMI and automatically cut off upon receiving a takeover request?

The automatic cessation of tasks by the HMI could theoretically reduce the time required for a driver to disengage from the NDRA and focus on the driving task. However, it is also possible that, even with automatic disengagement, participants may still require additional time to mentally transition from the task they were engaged in and back to the driving environment. This potential cognitive delay could negate some of the benefits of automatic task cessation and suggests that simply cutting off the NDRA may not be sufficient to ensure a safe and timely takeover. Research could investigate whether such automatic systems indeed lead to quicker and safer takeovers, or whether drivers still need additional cues or time to mentally adjust to the driving task after disengagement.

Training the UiC to optimise takeover performance

Investigating how expert drivers build SA could provide valuable insights into what constitutes an optimal response and how to effectively build situational awareness SA within the crucial 10-second window. Shaw et al. (2020) found that behavioural training, specifically using the CHAT (CHeck, Assess, and Takeover) checklist, significantly reduced the time drivers spent glancing back at non-driving related activities (NDRAs) after being notified to take over control in a level 3 automated vehicle. The behaviourally trained group reduced their NDRA glances to just 1.8 seconds, compared to 11.2 seconds for those trained with only an operating manual. This study highlights the critical role of targeted training in improving SA and safety during the transition from automated to manual driving.

Designing instructional materials such as a video that could be played in the HMI, could help standardise and improve the performance of everyday drivers. These materials could include guidelines on prioritising actions during a takeover, such as the importance of properly disengaging from NDRAs, conducting mirror checks, and quickly yet safely resuming control of the vehicle. Training programs that simulate various takeover scenarios could also be developed to reinforce these behaviours, ensuring that the UiC is better prepared to handle real-world situations.

4.5 Conclusion

This project explored the impact of various NDRAs on the SA and takeover performance of a UiC of self-driving vehicles. While some NDRAs may be safely performed, many can significantly impair SA and delay the transition to manual control, particularly in complex driving scenarios like roadworks. The project highlighted the need for refined mechanisms to measure SA and establish appropriate thresholds for safe takeovers and the importance of providing clear and specific instructions to drivers in automated vehicles to ensure that they understand how to conduct a safe takeover. This includes not just taking control quickly but doing so in a manner that ensures they are fully prepared to resume manual driving safely. The variability in participant responses and the influence of environmental factors suggest that further research is necessary to fully understand the nuances of NDRA impacts across different scenarios. This ongoing research will be crucial for developing informed policies and enhancing the safety of automated driving systems.

References

Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, *36*(7), 623–636. https://doi.org/10.1016/j.ergon.2006.04.002

BSI. (2020). Guidelines for developing and assessing control systems for automated vehicles. *BSI Standards*, *PAS 1880:2*.

Burnett, Gary, Large, David. R., Salanitri, D. (2019). *How will drivers interact with vehicles of the future?* RAC Foundation.

Cao, Z., Chuang, C. H., King, J. K., & Lin, C. T. (2019). Multi-channel EEG recordings during a sustained-attention driving task. *Scientific Data*, *6*(1), 19. https://doi.org/10.1038/s41597-019-0027-4

Centre for Connected & Autonomous Vehicles. (2021). *Safe Use of Automated Lane Keeping System (ALKS) Summary of Responses and Next Steps*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/980742/safe-use-of-automated-lane-keeping-system-alks-summary-of-responses-and-next-steps.pdf

Chen, H., Zhao, X., Li, Z., Li, H., Gong, J., & Wang, Q. (2023). Study on the influence factors of takeover behavior in automated driving based on survival analysis. *Transportation Research Part F: Traffic Psychology and Behaviour*, *95*, 281–296. https://doi.org/10.1016/j.trf.2023.04.012

Coster, X. A. (2015). *A Method for Determining How Much Time People Need to Understand Traffic Situations* [Delft University of Technology]. https://repository.tudelft.nl/islandora/object/uuid:4225c90b-9911-4c50-b32c-892315e4c440

Department for Transport. (2022). *Rules on the safe use of automated vehicles: summary of responses and government response*. https://www.gov.uk/government/consultations/safe-use-rules-for-automated-vehicles-av/outcome/rules-on-the-safe-use-of-automated-vehicles-summary-of-responses-and-government-response#draft-amendment-to-the-highway-code--a-new-section-for-self-driving-vehic

Department of Transport. (2023). *The Highway Code*. https://www.gov.uk/guidance/the-highway-code

Dogan, E., Honnêt, V., Masfrand, S., & Guillaume, A. (2019). Effects of non-driving-related tasks on takeover performance in different takeover situations in conditionally automated driving. *Transportation Research Part F: Psychology and Behaviour*, *62*, 494–504. https://doi.org/10.1016/j.trf.2019.02.010

Du, N., Zhou, F., Pulver, E., Tilbury, D., Robert, L., Pradhan, A., & Yang, J. (2020). Examining the Effects of Emotional Valence and Arousal on Takeover Performance in Conditionally Automated Driving. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3518015

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

Endsley, M. R. (1988b). Situation awareness global assessment technique (SAGAT). *Aerospace and Electronics Conference*, 789–795. https://doi.org/10.1109/NAECON.1988.195097

Endsley, M. R. (2015). Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 4–32. https://doi.org/10.1177/1555343415572631

Endsley, M. R. (2020). Situation awareness in driving. In *D. Fisher, W. J. Horrey, J. D. Lee & M. Regan (Eds.), Handbook of Human Factors for Automated, Connected and Intelligent Vehicles.* Taylor and Francis.

Endsley, M. R., Bolte, B., & Jones, D. . (2003). *Designing for Situation Awareness: An Approach to User-Centered Design* (N. Y. T. & Francis (ed.); 2nd ed.). CRC Press, Inc. Boca Raton, FL, USA. https://doi.org/10.1201/9780203485088

Endsley, M. R., & Jones, D. . (1996). Sources of situation awareness errors in aviation. *Aviation Space and Environmental Medicine*, *67*(6), 507–512. https://doi.org/10.1007/s12369-014-0266-7

Endsley, M. R., Selcon, S. J., Hardiman, T. D., & Croft, D. G. (1998). A comparative evaluation of SAGAT and SART for evaluations of SA. *The Human Factors and Ergonomics Society Annual Meeting*, *42*, 82–86.

Eriksson, A., & Stanton, N. A. (2016). Takeover Time in Highly Automated Vehicles : Noncritical Transitions to and From Manual Control. *Human Factors*, *59*(4), 689–705. https://doi.org/10.1177/0018720816685832

Eriksson, A., & Stanton, N. A. (2017). Driving Performance After Self-Regulated Control Transitions in Highly Automated Vehicles. *Human Factors*, *59*(8), 1233–1248. https://doi.org/10.1177/0018720817728774

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. https://doi.org/10.1007/978-3-662-53305-5_22

Gold, C., Berisha, I., & Bengler, K. (2015). Utilization of drivetime - Performing non-driving related tasks while driving highly automated. *Proceedings of the Human Factors and Ergonomics Society*, *2015-Janua*, 1666–1670.

https://doi.org/10.1177/1541931215591360

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

Gold, C., Körber, M., Lechner, D., & Bengler, K. (2016). Taking over Control from Highly Automated Vehicles in Complex Traffic Situations. *Human Factors*, *58*(4), 642–652. https://doi.org/10.1177/0018720816634226

Gugerty, L. J. (1997). Situation Awareness during Driving: Explicit and Implicit Knowledge in Dynamic Spatial Memory. *Journal of Experimental Psychology: Applied*, *3*(1), 42–66. https://doi.org/10.1037/1076-898X.3.1.42

Gugerty, L. J. (2011). Situation awareness in driving. *Handbook for Driving Simulation in Engineering, Medicine and Psychology*, 265–272. https://doi.org/10.1518/001872008X288394

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. *Proceedings of The Human Factors and Ergonomics Society 50th Annual Meeting*, 904–908.

Heo, J., Lee, H., Yoon, S., & Kim, K. (2022). Responses to Take-Over Request in Autonomous Vehicles: Effects of Environmental Conditions and Cues. *IEEE Transactions on Intelligent Transportation Systems*, *23*(12), 23573–23582. https://doi.org/10.1109/TITS.2022.3201074

Jenness, J. W., Lattanzio, R. J., Toole, M. O., & Taylor, N. (2002). *Voice-Activated Dialing Or Eating A Cheeseburger : Which Is More Distracting During Simulated Driving ?* 592–596.

Jiang, T., Wang, Y., & Tang, R. (2024). Playing Games Guiding Attention Improves Situation Awareness and Takeover Quality during Automated Driving. *International Journal of Human-Computer Interaction*, *40*(8), 1892–1905. https://doi.org/10.1080/10447318.2023.2228068

Jones, D., & Endsley, M. R. (1996). Sources of Situation Awareness Errors in Aviaiton Domain. *Aviation Space and Environmental Medicine*, *67*(6), 507–512. https://doi.org/10.1039/c4qo00187g

Kastle, J. L., Anvari, B., Krol, J., & Wurdemann, H. A. (2021). *Neurocomputing Correlation between Situational Awareness and EEG signals*. *432*, 70–79. https://doi.org/10.1016/j.neucom.2020.12.026

Kunze, A., Summerskill, S. J., Marshall, R., & Filtness, A. J. (2019). Automation transparency: implications of uncertainty communication for human-automation interaction and interfaces. *Ergonomics*, *62*(3), 345–360. https://doi.org/10.1080/00140139.2018.1547842

Large, D. ., Burnett, G., & Salanitri, D. (2019). *A Longitudinal Simulator Study to Explore*

Drivers ' Behaviour in Level 3 Automated Vehicles. 222–232. https://doi.org/10.1145/3342197.3344519

Large, D. R., Burnett, G. E., Andrew, M., Muthumani, A., & Matthias, R. (2019). A longitudinal simulator study to explore drivers ' behaviour during highly- automated driving. *Figshare.*, 0–10. https://hdl.handle.net/2134/26238

Law Commission of England and Wales & Scottish Law Commission. (2022a). *Automated Vehicles: joint report* (Issue 404).

Law Commission of England and Wales & Scottish Law Commission. (2022b). *Automated Vehicles: Summary of joint report*. *January*(404).

Li, M., Feng, Z., Zhang, W., Wang, L., Wei, L., & Wang, C. (2023). How much situation awareness does the driver have when driving autonomously? A study based on driver attention allocation. *Transportation Research Part C: Emerging Technologies*, *156*. https://doi.org/10.1016/j.trc.2023.104324

Liang, N., Yang, J., Yu, D., Prakah-Asante, K. O., Curry, R., Blommer, M., Swaminathan, R., & Pitts, B. J. (2021). Using eye-tracking to investigate the effects of pre-takeover visual engagement on situation awareness during automated driving. *Accident Analysis and Prevention*, *157*(January), 106143. https://doi.org/10.1016/j.aap.2021.106143

Lo, J. C., Sehic, E., Brookhuis, K. A., & Meijer, S. A. (2016). Explicit or implicit situation awareness? Measuring the situation awareness of train traffic controllers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *43*, 325–338. https://doi.org/10.1016/j.trf.2016.09.006

Lorenz, L., Kerschbaum, P., & Schumann, J. (2014). Designing take over scenarios for automated driving: How does augmented reality support the driver to get back into the loop? *Proceedings of the Human Factors and Ergonomics Society*, *2014-January*, 1681–1685. https://doi.org/10.1177/1541931214581351

Louw, T., Merat, N., & Jamson, H. (2015). Engaging with highly automated driving. To be or not to be in the loop. *8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Salt Lake City, Utah, USA*, *1996*, 190–196. https://doi.org/10.13140/RG.2.1.2788.9760

Lu, Z., Coster, X., & de Winter, J. (2017). How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving. *Applied Ergonomics*, *60*, 293–304. https://doi.org/10.1016/j.apergo.2016.12.003

Marti, P., Jallais, C., Koustanaï, A., & Guillaume, A. (2022). Impact of the driver's visual engagement on situation awareness and takeover quality. *Transportation Research Part F: Psychology and Behaviour*, *87*(September 2021), 391–402. https://doi.org/10.1016/j.trf.2022.04.018

Melcher, V., Rauh, S., Diederichs, F., Widlroither, H., & Bauer, W. (2015). Take-Over

Requests for Automated Driving. *Procedia Manufacturing*, *3*(Ahfe), 2867–2873. https://doi.org/10.1016/j.promfg.2015.07.788

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

Mok, B., Johns, M., Miller, D., & Ju, W. (2017). Tunneled In: Drivers with Active Secondary Tasks Need More Time to Transition from Automation. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, 2840–2844. https://doi.org/10.1145/3025453.3025713

Mutzenich, C., Durant, S., Helman, S., & Dalton, P. (2021). Situation Awareness in Remote Operators of Autonomous Vehicles: Developing a Taxonomy of Situation Awareness in Video-Relays of Driving Scenes. *Frontiers in Psychology*, *12*(November). https://doi.org/10.3389/fpsyg.2021.727500

Niklasson, L., Riveiro, M., Johansson, F., Dahlbom, A., Falkman, G., Ziemke, T., Brax, C., & Kronhamn, T. (2007). A Unified Situation Analysis Model for Human and Machine Situation Awareness. *Herzog, O., Rödiger, K.-H., Ronthaler, M. & Koschke, R. (Hrsg.), Informatik 2007 – Informatik Trifft Logistik – Band 2. Bonn: Gesellschaft Für Informatik e. V.. (S. 105-109).*, *2*, 105–109.

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

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. *Proceedings of the Human Factors and Ergonomics Society*, *2014-Janua*(1988), 2063–2067. https://doi.org/10.1177/1541931214581434

SAE International. (2018). Surface Vehicle Recommended Practice. In *J3016*.

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, *39*(3), 490–500. https://doi.org/10.1016/j.ergon.2008.10.010

Salmon, P. M., Stanton, N. A., & Young, K. L. (2012). Situation awareness on the road: Review, theoretical and methodological issues, and future directions. *Theoretical Issues in Ergonomics Science*, *13*(4), 472–492. https://doi.org/10.1080/1463922X.2010.539289

Shariati, A., Anvari, B., K, J. L., Krol, J., Stanton, N. A., & Reed, N. (2023). *Analysing Variables Associated with Driver Reaction during the Transition from Automated to Manual Driving*. *3*.

Shaw, E., Large, D. R., Burnett, G., & Foundation, R. A. C. (2020). *Driver training for future automated vehicles: introducing CHAT (CHeck, Assess and Takeover)*. *November*, 113p. https://www.racfoundation.org/research/safety/driver-training-for-future-automated-vehicles%0Ahttps://trid.trb.org/view/1756392

Sun, X., Cao, S., & Tang, P. (2021). Shaping driver-vehicle interaction in autonomous vehicles: How the new in-vehicle systems match the human needs. *Applied Ergonomics*, *90*. https://doi.org/10.1016/j.apergo.2020.103238

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. https://doi.org/10.1038/381520a0

United Nations. (2021). Uniform provisions concerning the approval of vehicles with regard to Automated Lane Keeping Systems. *UN Regulation No. 157*.

Van Miltenburg, M. M. P. G., Lemmers, D. J. A., Tinga, A., Christoph, M., & Zon, R. (2022). Can EEG Measurements be Used to Estimate the Performance of Taking over Control from an Autonomous Vehicle for Different Levels of Distracted Driving? An Explorative Study. *Adjunct Proceedings - 14th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2022*, 20–24. https://doi.org/10.1145/3544999.3552324

Vogelpohl, T., Kühn, M., Hummel, T., Gehlert, T., & Vollrath, M. (2018). Transitioning to manual driving requires additional time after automation deactivation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *55*, 464–482. https://doi.org/10.1016/j.trf.2018.03.019

Wulf, F., Rimini-Doring, M., Arnon, M., & Gauterin, F. (2015). Recommendations Supporting Situation Awareness in Partially Automated Driver Assistance Systems. *IEEE Transactions on Intelligent Transportation Systems*, *16*(4), 2290–2296. https://doi.org/10.1109/TITS.2014.2376572

Yang, L., Ma, R., Zhang, H. M., Guan, W., & Jiang, S. (2018). Driving behavior recognition using EEG data from a simulated car-following experiment. *Accident Analysis and Prevention*, *116*(November 2017), 30–40. https://doi.org/10.1016/j.aap.2017.11.010

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 and Prevention*, *92*, 230–239. https://doi.org/10.1016/j.aap.2016.04.002

Acknowledgements[1]

The Loughborough team contributed significantly to the project throughout all phases. Firstly, the team contributed to the development of the Situational Awareness (SA) metric, helping to devise innovative methods to measure SA, using decades of expertise related to human factors in driving. Secondly, Loughborough took the lead on scenario design, defining the specific parameters required to achieve the study aims, subsequently creating bespoke scenarios in the SCANeR environment for the project. As Loughborough were the first data collection site, they were instrumental in developing and refining the trial protocols, test plans and ethical clearance processes appropriate for the project as a whole. The team then conducted live testing with 31 participants, keeping a detailed log of the experiments, and responding agilely to any required changes and developments. Following data collection, the Loughborough team processed data from the vehicle sensors, eye tracking glasses and GoPro recordings, to provide data in an accessible format for analysis. Finally, the Loughborough team contributed significantly to defining the analysis parameters that were required to inform the SA metric, in particular with respect to the human behaviour measures, as well as contributing to the final report by providing expertise in interpreting the results.

UCL's team have contributed significantly to the project by drafting scenarios, developing innovative methods to measure Situational Awareness (SA), and constructing scenarios in Unity for the IM@UCL driving simulator, all aligned with DfT requirements. The team conducted a thorough literature review, which informed both the scenario drafting and the method development. UCL also successfully prepared and integrated new hardware and sensors, secured ethical approval, and fine-tuned experimental details during piloting. They played a key role in conducting pilot testing, live testing with 66 participants, and ensuring data management adhered to GDPR standards. UCL expertly processed data from vehicular sensors, eye-tracking, and GoPro recordings, transforming it into meaningful KPIs and providing insights into gaze and physiological patterns. Additionally, they supported the ranking of KPIs and offered valuable advice on statistical analysis. Although UCL's role did not extend to final parameter selection, statistical analysis, or report writing, their contributions, including the literature review, were instrumental to the success of the project.

Conceptualization: UCL & LU team; **Methodology**: UCL & LU team; **Investigation**: UCL & LU team; **Software**: UCL & LU team; **Data Curation**: UCL & LU team; **Formal Analysis**: UCL & LU team (advisory role); **Project Administration**: UCL & LU team; **Resources**: UCL & LU team; **Writing – Review & Editing**: UCL & LU team (technical details regarding IM@UCL & LU driving simulators)

[1] CRediT

5. Appendix

5.1 Disengagement coding decisions

Non-Driving Related Activity (NDRA)	Full Engagement	Partial Disengagement
Cradled mobile	Stopping the film on the cradled handset by pressing pause.	Attempting to stop the film and failing OR leaving the film playing and taking over manual control OR returning back to the cradled handset after the TO.
Handheld mobile	Pressing pause on the game and placing the phone on the passenger seat.	Placing the phone on the passenger seat without pausing the game (sound continues playing) OR dropping the phone onto lap and taking over manual driving.
Magazine	Placing the magazine on the passenger seat.	Starting driving holding the magazine OR dropping the magazine onto lap.
Wordsearch	Placing the book and the pen on the passenger seat.	Starting driving holding the pen OR dropping the wordsearch book onto lap OR holding the wordsearch book against the steering wheel and starting driving
Drinking water	Placing the cup in the cup holder to the left of the driver (centre console).	Starting driving holding the cup.
Eating popcorn	Placing the bag of popcorn on the passenger seat (holster remains around the participant’s neck as not possible to remove).	Dropping the bag into lap and started driving OR holding the bag against the steering wheel and starting driving.

5.2 Specialist group

We express our sincere gratitude to the specialist group whose expertise and collaboration were extremely helpful to this project. These individuals contributed to the review of the NDRA design and participated in a workshop to evaluate results and discuss future research proposals. Their collective efforts and insights have been instrumental in advancing this work. The individuals involved were:

- Andy Cumming and Chrissie Hare from Jaguar Land Rover
- Professor Stewart Birrell from Coventry University
- Cyriel Dials from the Royal College of Art's Intelligent Mobility Design Centre
- Melissa Gilbert from the Motability Foundation
- Catherine Bowen and Amanda Brandon from the BVRLA

5.3 Sample breakdown

Demographic	Criteria	UCL	LU
Age	18-25	15	7
	26-35	12	3
	36-45	14	7
	46-65	13	6
	66+	9	1
Gender	Male	37	10
	Female	26	14
	Other	0	0
Ethnicity	White	26	17
	Mixed or Multiple ethnic groups	4	3
	Asian or Asian British	15	1
	Black, Black British, Caribbean or African	8	2
	Other	10	1
Socio-economic group	B	11	0
	C1	37	12
	C2	15	11
	D	0	1
	E	0	0
Type of location	Large city	46	3
	Small city	4	1
	Town	12	7
	Village	0	7
	Rural	1	6
Amount of Years of Driving Experience	0 to 1 Year	8	2
	2 to 5 Years	8	5
	6 to 10 Years	11	2
	11 to 15 Years	6	3
	16+ Years	30	12
Main Reason for Driving	Commuting	32	16
	Business	14	5
	Leisure	54	20
	Errands	37	17
	School Pick-up drop off	11	2
How frequently they drive	Everyday	32	18
	Most Days	13	2
	2-3 times a week	10	3
	Once a week	1	0
	1-2 times a month	3	0
	Once a month	1	0
	Rarely	2	0
	Never	1	1
Attitude towards tech	Positive	54	20
	Negative	9	4

5.4 Notes and rationale for analysis approach

Statistical analysis rationale

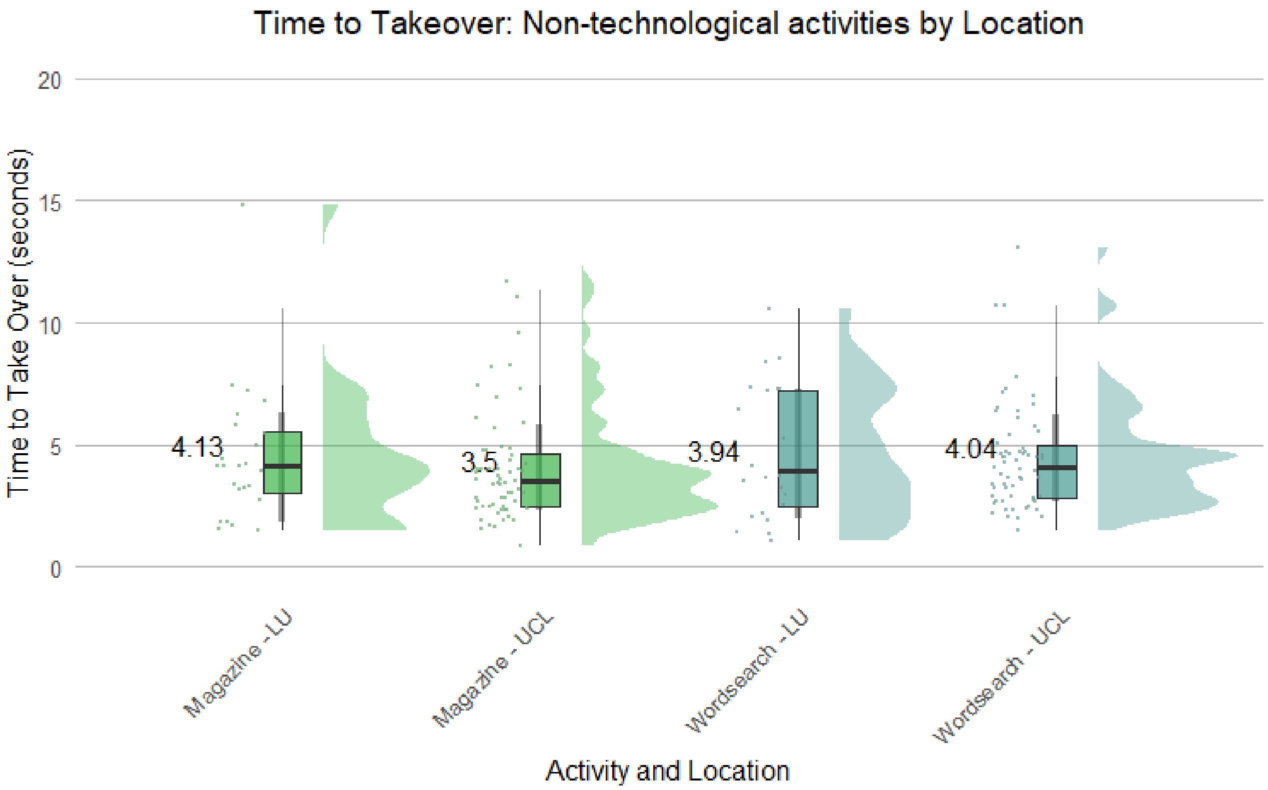
Linear Mixed Models (LMM) have been used for statistical analysis. This approach allowed us to focus on individual performance across different scenarios and activities, rather than relying on average performance metrics, acknowledging significant variations in participant performance. For example, some participants may respond faster than others to takeover requests or may improve over time, or some activities may be perceived as more difficult than others. Additionally, in this study, some eye-tracking data was missing, and some participants took longer than 30 seconds or did not take over at all, leading to their data being excluded. Furthermore, some disengagements were not recorded due to missing GoPro data for those trials due to technical failures. LMM allows for missing data, whereas traditional repeated measures ANOVAs do not, providing a clear advantage of LMM analysis over ANOVA for this study.

In LMM analysis, one category from each variable (like Congestion for Scenario Type and Handheld mobile for NDRA) is used as the reference or baseline. The other categories are compared to this baseline to show how they differ in their effects. The reference isn't shown in the results, but all the estimates are relative to it. The fixed effects, random effects and model fit statistics output of all LMM tests are in the Technical Appendix (Section 6).

To provide a clearer picture for readers, we report mean times in visualisations in Section 3, giving an intuitive understanding of the time to take over in seconds. This helps to easily grasp how long participants typically take to regain control of the vehicle. However, for statistical analysis, we consider participant variables as random effects, which means we can account for individual differences among participants such as being faster or slower to take over or poor simulator control. This approach helps us to understand how different activities affect the time it takes for someone to take control of the vehicle. By focusing on the differences caused by the activities rather than differences between the individuals, we can get a clearer picture of the true impact of those activities on takeover times. While the mean times provide a straightforward summary, the detailed statistical analysis using LMM offers a more precise understanding of the underlying patterns and effects. Full details of LMM analyses are available in the Technical Appendix/Report.

Visualisations

Bar charts and raincloud plots are used to visualise the data across NDRAs, scenarios and locations. In this report we use both to communicate a visual representation of the data. A raincloud plot is a powerful and easy-to-understand way to show a lot of data at once. It combines several types of charts into one, helping you see the overall trends and the individual details. An example plot is shown below:



Below is a tutorial of how to read the data:

- Box Plot: A way to visually display the distribution of data. In the middle of the box, there's a line that represents the median, which is the middle value when you arrange all your data points from lowest to highest. The box itself shows where most of the data lies. The top and bottom edges of the box represent the range where the middle 50% of the data falls, so it gives a sense of where most of the results are grouped. In other words, it helps you see the typical range and the central tendency of the data at a glance. Additionally, the box itself typically shows the interquartile range (IQR), which represents the middle 50% of the data, while the "whiskers" extend to the minimum and maximum values within 1.5 times the IQR from the quartiles, excluding outliers.
- Cloud (Density Plot): Shows where most of the data points are concentrated. On the right side, you'll see a shape that looks like a cloud. The thicker or denser the cloud in a particular area, the more data points are around that value. This helps you quickly see where most of the data is grouped and how it spreads out across different values. It's a visual way to understand the distribution and concentration of your data.
- Raindrops (Raw Data Points): To the left of the box plot, you see individual dots, like raindrops. Each dot represents a single person's result. This helps you see every single data point clearly.
- Jittered Points: Sometimes, the dots (raindrops) are spread out horizontally, so they don't overlap too much if many participants have the same score. This makes it easier to see each one.

5.5 Pre and Post study questionnaire

Pre- study questionnaire

Q1. Did you purchase any of the following products / subscribe to any of the following services over the last two years? Tick all that apply

- Smartphone (iPhone, Android or other)
- Cloud (paid subscription, not free account) [iCloud excluded]
- Home assistant (Google Home, Amazon Echo or Apple HomePod)
- Mobile Payment (ApplePay, GooglePay, SamsungPay or others)
- Home Internet of Things (IoT) devices (i.e. connected home devices like Wi-Fi cameras, connected appliances, smart doorbell (Ring), Amazon Dash buttons, etc.)
- High end Tech accessories (VR Headsets, 360°camera) [GoPro excluded]
- Smartwatches (Android, Apple, Garmin, Polar)
- Gaming hardware (VR Headsets, Gaming specific PC, consoles)
- None of the above

Q2. How comfortable are you with accepting the following transport innovations in the future? Please rate each on a scale of 1 to 5, where 1 indicates "Not comfortable at all" and 5 indicates "Very comfortable."

- Electric vehicles (EVs)
- Automated vehicles (AVs)
- Ride-sharing services (e.g., Uber, Lyft)
- Bike-sharing programmes
- Electric scooters
- Electric bicycles (e-bikes)
- Augmented reality (AR) navigation for pedestrians
- Automated Lane Keeping Systems (ALKS)
- Parking Assist
- Adaptive cruise control

Q3. Do you hold any of the following UK driving licenses? Tick Yes or No for each option

- Motorbike
- Car (Automatic)
- Car (Manual)
- Large Vehicles (e.g., lorries)

Q4. How many years of driving experience do you have since passing your test? Tick how many years of driving experience you have for each of the options.

- Motorbike
- Car
- Large Vehicles

Q5. How often do you drive? Tick one option

- Everyday
- Most days
- 2-3 times a week
- Once a week
- 1 – 2 times a month
- Never

Q6. What is the purpose of your driving? Tick as many that apply.

- Commuting
- Business
- Leisure
- Errands
- School pick up/drop off
- Other: please specify

Q7. First of all, what is your current employment status? Tick one option

- Full-time employment
- Part-time employment
- Unemployed
- Freelance / Self-employed
- Homemaker
- Student
- Retired

Q8. Do you suffer with any of the following? Tick as many that apply

- Epilepsy
- Seizures
- Negative reactions to bright or flashing lights
- Motion sickness
- Sensitivity to sunlight (photosensitivity)
- Other neurological disorder, please specify.

Q9. What is your ethnicity? Tick one option

- White (includes English, Welsh, Scottish, British, Northern Irish, Gypsy, Irish Traveller, Roma or any other White background)
- Mixed or Multiple ethnic groups (include White and Black Caribbean, White and Black African, White and Asian or any other Mixed or multiple ethnic background)
- Asian or Asian British (include Indian, Pakistani, Bangladeshi, Chinese or any other Asian background)
- Black, Black British, Caribbean or African (includes Caribbean, African or any other Black, Black British or Caribbean background)

- Other ethnic group (includes Arab and any other ethnic group, please specify)
- Prefer not to say

Q10. Which gender do you identify as? Tick one option

INTERVIEWER: PLEASE SELECT ONE ANSWER.
- Male
- Female
- Trans male
- Trans female
- Non-binary
- Other gender identity, please specify.
- Prefer not to say.

Q11. How old are you? Write answer below.

Write in years: ______________________________

Post-study questionnaire

Q1. How confident were you in your situational awareness when taking back control from the CAV system? Please rate each on a scale of 1 to 5, where 1 indicates "Not comfortable at all" and 5 indicates "Very comfortable".

- 1
- 2
- 3
- 4
- 5

Q2. Were there any specific moments during the experiment where you felt more or less aware of your surroundings? Be specific.

Q3. How did you perceive the takeover request alerts during the experiment? Were they clear and easy to understand? Select Yes or No. If you select no explain why.
- Yes
- No

Q4. Did you feel prepared to take over control of the vehicle when prompted? Select Yes or No. If you select no explain why.
- Yes
- No

Q5. Did you have any problems understanding what you had to do when performing the non-driving related activities? Select Yes or No. If you select yes explain why.
- Yes

- No

Q6. Which non-driving related activities did you find more or less immersive while operating the CAV system? Please rate each on a scale of 1 to 5, where 1 indicates "Not immersed at all" and 5 indicates "Completely immersed, I forgot where I was"

- Drinking water
- Eating popcorn
- Phone handheld (playing Tetris)
- Phone cradle (watching a movie)
- Reading
- Completing a wordsearch/sudoku

Q7. Do you think any of the non-driving related activities should not be allowed while using the CAV system? Select all that apply.

- Drinking water
- Eating popcorn
- Phone handheld (playing Tetris)
- Phone cradle (watching a movie)
- Reading
- Completing a wordsearch/sudoku
- None of the above.

Q8. Were there any aspects of the experiment setup or procedures that you found particularly enjoyable or challenging? Be specific.

5.6 NASA-TLX & SART Questionnaires

NASA -TLX

Q1. How much mental activity was required e.g., thinking, deciding, looking. Was the task easy or demanding? (Mental demand)

Drag the slider to a point on the scale one end is Low and the other end is High.

Q2. How much physical activity was required (e.g., pushing, pulling, turning, controlling)? Was the task easy or demanding? (Physical demand)

Drag the slider to a point on the scale one end is Low and the other end is High.

Q3. How much time pressure did you feel due to the rate or pace at which the task occurred? Was the pace slow or frantic? (Temporal demand)

Drag the slider to a point on the scale one end is Low and the other end is High.

Q4. How hard did you have to work (mentally and physically) to accomplish your level of performance? (Effort)

Drag the slider to a point on the scale one end is Low and the other end is High.

Q5. How successful do you think you were in accomplishing the goals of the task set by the experimenter? How satisfied were you with your performance? (Performance)

Drag the slider to a point on the scale one end is Good and the other end is Poor.

Q6. How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task? (Frustration)

Drag the slider to a point on the scale one end is Low and the other end is High.

SART

Q1. How changeable was the situation (instability)

Drag the slider to a point on the scale 1-5. 1 is stable and straightforward and 5 is changing suddenly.

Q2. How many variables are changing within the situation (variability)?

Drag the slider to a point on the scale 1-5. 1 is very few variables changing and straightforward and 5 is a large numbers of factors varying.

Q3. How complicated is the situation (complexity)?

Drag the slider to a point on the scale 1-5. 1 is stable and straightforward and 5 is complex with many inter-related components.

Q4. How aroused are you in the situation (arousal)?

Drag the slider to a point on the scale 1-5. 1 is a low degree of alertness and straightforward and 5 is alert and ready for activity.

Q5. How much mental capacity do you have to spare in the situation (spare capacity)?

Drag the slider to a point on the scale 1-5. 1 is nothing to spare at all and 5 is sufficient to attend many variables.

Q6. How much are you concentrating on the situation (concentration)?

Drag the slider to a point on the scale 1-5. 1 is focusing only on one aspect and 5 is concentrating on many aspects of the situation.

Q7. How low is your attention divided in the situation (attention division)?

Drag the slider to a point on the scale 1-5. 1 is focusing only on one aspect and 5 is concentrating on many aspects of the situation.

Q8. How much information have you gained about the situation (quantity)?

Drag the slider to a point on the scale. One point is very little, and the other end of the scale is a great deal of knowledge.

Q9. How familiar are you with the situation (familiarity)?

Drag the slider to a point on the scale. Lowest being this is a new situation, and the other end of the scale is a great deal of relevant experience.

5.7 Eye tracking analysis pipeline

Eye tracking analysis pipeline involves several key steps, leveraging eye-tracking technology to capture and analyse participants' visual attention. The primary focus is on understanding how participants allocate their gaze across various AOIs during different activities and scenarios. The detailed steps of the analysis pipeline are as follows:

Identification of AOIs:
Key AOIs were identified that participants should focus on for each activity and scenario to measure SA. These included the road, rear-view mirror, right-side mirror, speedometer, HMI, NDRAs and other areas of the car (e.g., the centre console where the water cup was placed).

Data Collection:
Eye-tracking glasses were used to record the gaze points of participants in real-time during the driving simulator sessions. Single frames from these recordings were analysed to ensure the accurate detection of AOIs.

Gaze Point Overlay:
Detected AOIs were overlaid with participants' gaze points to visually represent where and how long participants looked at each AOI during the activities. This visual overlay helps in identifying the focus areas and understanding the participants' attention distribution.

Retention of Bright Large Areas:
Bright and large areas within the gaze recordings were retained to highlight the prominent AOIs. This step ensures that key focus areas such as the road and mirrors are clearly visible in the analysis.

Fixation Analysis:
The frequency and duration of fixations on each AOI were calculated. This analysis focused on how often and for how long participants looked at the road,

mirrors, speedometer, and HMI. Fixations on "Other areas" (areas not predefined as AOIs) were also tracked to identify if participants were searching for additional information, indicating potential gaps in situational awareness.

Scenario and Task Variations:

The analysis compared the number of fixations and the duration of looking times across different activities and scenarios. This step was crucial in understanding how activity complexity and scenario type influenced visual attention and situational awareness.

6. Technical Appendix

6.1 Disengagement LMM Results

A binomial model was utilised to investigate the effect of scenario type and activity on disengagement. In this model, scenario type and activity were treated as fixed effects, while location and participant (nested within location) were treated as random effects to account for variability across both locations and individuals. Given that disengagement is a binary outcome (Yes/No), a logit link function was used. The "bobyqa" optimiser was applied to ensure model stability and accurate convergence. This approach accounts for participant differences and allows for a more robust understanding of how scenario type and activity influence disengagement. Full model details can be found in Technical Appendix 6.1.

Fixed Effects:

Model Formula: Disengage ~ Scenario Type + Activity + (1 | Location/Participant)					
Effect	**Estimate**	**Standard Error**	**z-value**	**p-value**	**Significance**
(Intercept)	-0.571	0.987	-0.578	0.563	
Scenario Type Roadworks	-0.184	0.133	-1.383	0.167	
Cradled mobile phone	-0.919	0.196	-4.685	0.000	***
Magazine	5.575	0.274	20.338	0.000	***
Wordsearch	4.010	0.224	17.875	0.000	***
Drinking water	4.965	0.251	19.786	0.000	***

Model Formula: Disengage ~ Scenario Type + Activity + (1 | Location/Participant)					
Effect	**Estimate**	**Standard Error**	**z-value**	**p-value**	**Significance**
Eating popcorn	3.962	0.223	17.753	0.000	***

Significance. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant:Location	(Intercept)	6.520	2.553
Location	(Intercept)	1.699	1.303

Model Fit Statistics:

Statistic	Value
AIC	2,107.816
BIC	2,163.188
Log Likelihood	-1,044.908
Deviance	1,762.110

6.2 Disengagement by Location LMM Results

A binomial model was employed to examine the relationship between disengagement, activity, and location. The model treated activity as a fixed effect and participant as a random effect, nested within location, to account for potential variability between individuals. Given the binary nature of the disengagement outcome (Yes/No), a logit link function was used within the binomial model. The "bobyqa" optimiser was applied to ensure stable convergence of the model. This approach helps address the small sample size and participant variability, particularly at LU (n=24), allowing for more robust statistical analysis of the disengagement patterns.

Fixed Effects:

Model Formula: Disengage ~ Location + Activity + (1 | Participant)					
Effect	**Estimate**	**Standard Error**	**z-value**	**p-value**	**Significance**
(Intercept)	0.800	0.585	1.368	0.171	
Location: UCL	-2.784	0.666	-4.183	0.000	***
Cradled mobile phone	-0.905	0.197	-4.585	0.000	***
Magazine	5.575	0.275	20.247	0.000	***
Wordsearch	4.008	0.225	17.807	0.000	***
Drinking water	4.956	0.252	19.703	0.000	***
Eating popcorn	3.959	0.223	17.724	0.000	***

Significance. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Random Effects:

Group	**Random Effect**	**Variance**	**Standard Deviation**
Participant	(Intercept)	6.375	2.525

Model Fit Statistics:

Statistic	**Value**
AIC	2,100.814
BIC	2,150.034
Log Likelihood	-1,042.407
Deviance	1,763.727

6.3 Disengagement by Scenario LMM Results

Fixed Effects:

Model Formula: Disengage ~ Scenario Type + (1 | Location/Participant)					
Effect	Estimate	Standard Error	z-value	p-value	Significance.
(Intercept)	1.317	0.515	2.560	0.01	**
Scenario Type Roadworks	-0.099	0.081	-1.226	0.22	

Significance. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Random Effects:

Group	Random Effect	Variance	Std. Deviation
Participant:Location	(Intercept)	1.762	1.328
Location	(Intercept)	0.464	0.681

Model Fit Statistics:

Statistic	Value
AIC	3,817.314
BIC	3,841.924
Log Likelihood	-1,904.657
Deviance	3,522.092

6.4 Time to Takeover by Activity and Scenario LMM Results

A nested random intercept model was used to examine the effects of activity and scenario type on reaction times, with location and participant considered as nested random effects. In this model, activity and scenario type were treated as fixed effects, while participant was nested within location to account for the variability both between locations and between participants within each location. The dependent variable, reaction time (rt_log), was log-transformed due to its non-normal distribution, allowing for a more accurate model fit. This nested structure helps capture the influence of both individual and location-specific factors on reaction times. The smaller sample size at LU (n=24) may have contributed to the increased variability observed in the data, which should be taken into account when interpreting these results.

Fixed Effects:

Model Formula: rt log ~ Activity + Scenario Type + (1 | Location/Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
(Intercept)	1.266	0.070	1.408	18.036	0.012	*
Drinking water	-0.244	0.057	596.419	-4.304	0.000	***
Eating popcorn	-0.056	0.057	596.205	-0.988	0.323	
Handheld mobile phone	0.078	0.057	596.419	1.376	0.169	
Magazine	0.015	0.057	596.324	0.257	0.797	
No NDRA (Congestion)	-0.300	0.059	596.342	-5.076	0.000	***
No NDRA (Roadworks)	-0.320	0.058	596.287	-5.439	0.000	***
Wordsearch	0.088	0.057	596.324	1.562	0.119	
Scenario Type Roadworks	0.065	0.033	596.202	2.002	0.046	*

Significance. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant:Location	(Intercept)	0.187	0.432
Location	(Intercept)	0.002	0.039
Residual		0.138	0.371

Model Fit Statistics:

Statistic	Value
AIC	859.624
BIC	914.081
Log Likelihood	-417.812
Deviance	835.624

6.5 Interaction between Location and Activity LMM Results

An interaction model was employed to explore the combined effects of activity and location on reaction times, with scenario type also included as a fixed effect. This model tests for an interaction between activity and location, meaning it examines whether the effect of the activity on reaction time differs depending on the location. Additionally, participant is treated as a random effect to account for individual variability. The dependent variable, reaction time (rt_log), was log-transformed to address its non-normal distribution. By incorporating the interaction term, this model provides a more nuanced understanding of how activity and location jointly influence reaction times.

Fixed Effects:

Model Formula: rt log ~ Activity * Location + Scenario Type + (1 | Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
(Intercept)	1.343	0.117	212.272	11.434	0.000	***
Drinking water	-0.169	0.107	589.107	-1.572	0.117	
Eating popcorn	-0.011	0.107	589.107	-0.101	0.920	
Handheld mobile phone	0.076	0.107	589.107	0.711	0.477	
Magazine	-0.026	0.107	589.107	-0.239	0.811	
No NDRA (Congestion)	-0.276	0.108	589.109	-2.542	0.011	*
No NDRA (Roadworks)	-0.279	0.108	589.109	-2.574	0.010	**
Wordsearch	-0.014	0.107	589.107	-0.129	0.897	
Location:UCL	-0.114	0.137	206.727	-0.834	0.405	
Scenario Type Roadworks	0.066	0.033	589.201	2.011	0.045	*

Model Formula: rt log ~ Activity * Location + Scenario Type + (1 | Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
Drinking water: Location:UCL	-0.105	0.1265312	589.230	-0.832	0.406	
Eating popcorn: Location:UCL	-0.063	0.1264867	589.143	-0.496	0.620	
Handheld mobile phone: Location:UCL	0.003	0.1265383	589.228	0.021	0.983	
Magazine: Location:UCL	0.056	0.1263652	589.191	0.439	0.661	
No NDRA (Congestion): Location:UCL	-0.033	0.1263632	589.191	-0.262	0.794	
No NDRA (Roadworks): Location:UCL	-0.056	0.1263632	589.191	-0.443	0.658	
Wordsearch: Location:UCL	0.141	0.126	589.191	1.118	0.264	

Significance. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant	(Intercept)	0.186	0.432
Residual		0.138	0.372

Model Fit Statistics:

Statistic	Value
AIC	888.618
BIC	974.843
Log Likelihood	-425.309
Deviance	850.618

6.6 Time to Target Speed - Congestion LMM Results

The model evaluates the effect of NDRA and whether the participant reached target speed and the time taken to reach target speed during the congestion scenario. Location and participant are treated as nested random effects to account for variability between locations and individual differences. This model provides insight into how activity influences time to reach target speed in a congested environment while controlling for location-based differences.

Fixed Effects:

Model Formula: Time To Target Speed ~ Activity + Reached Target Speed + (1 | Location/Participant)

Effect	Estimate	Standard Error	df	t value	Pr(>|t|)	Significance
(Intercept)	5.006	1.266	1.007	3.954	0.156	
Cradled mobile phone	1.456	0.235	2,245.130	6.190	0.000	***
Drinking water	-0.466	0.233	2,235.991	-2.001	0.046	*
Eating popcorn	0.215	0.219	2,230.467	0.986	0.324	
Handheld mobile phone	2.542	0.225	2,232.339	11.306	0.000	***
Magazine	0.163	0.227	2,244.486	0.718	0.473	
Wordsearch	0.272	0.227	2,245.282	1.200	0.230	

Significance. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant:Location	(Intercept)	8.321	2.885
Location	(Intercept)	2.924	1.710
Residual		8.800	2.967

Model Fit Statistics:

Statistic	Value
AIC	11,846.044
BIC	11,903.463
Log Likelihood	-5,913.022
Deviance	11,826.044

6.7 Time to Target Speed - Roadworks LMM Results

This model assesses the effect of activity and reaching target speed on the time to target speed within the roadworks scenario. Again, location and participant are treated as nested random effects. This model helps to understand how activity impacts the ability to reach target speed in a roadworks setting, accounting for variability across locations and participants.

Fixed Effects:

Model Formula: Time To Target Speed ~ Activity + Reached Target Speed + (1 | Location/Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
(Intercept)	11.671	7.008	1.001194	1.665	0.344	
Cradled mobile phone	-0.965	0.376	1,704.689	-2.568	0.010	**
Drinking water	-0.890	0.377	1,701.399	-2.359	0.018	*
Eating popcorn	-3.075	0.385	1,697.438	-7.979	0.000	***
Handheld mobile phone	-0.064	0.362	1,712.520	-0.176	0.860	
Magazine	-2.105	0.388	1,722.823	-5.429	0.000	***
Wordsearch	-1.322	0.383	1,715.753	-3.453	0.001	***

Significance. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant:Location	(Intercept)	22.024	4.693
Location	(Intercept)	97.470	9.873
Residual		17.184	4.145

Model Fit Statistics:

Statistic	Value
AIC	10,263.145
BIC	10,317.859
Log Likelihood	-5,121.572
Deviance	10,243.145

6.8 Time to Target Speed by Location LMM Results

This model formula assesses the impact of activity, whether the participant reached target speed, and location on the time to target speed.

Fixed Effects:

Model Formula: Time To Target Speed ~ Activity + Reached Target Speed + Location + (1 | Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
(Intercept)	4.646	0.994	81.212	4.676	0.000	***
Cradled mobile phone	-0.965	0.376	1,704.691	-2.568	0.010	**
Drinking water	-0.890	0.377	1,701.399	-2.358	0.018	*
Eating popcorn	-3.076	0.385	1,697.435	-7.981	0.000	***

Model Formula: Time To Target Speed ~ Activity + Reached Target Speed + Location + (1 | Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
Handheld mobile phone	-0.064	0.362	1,712.520	-0.177	0.860	
Magazine	-2.105	0.388	1,722.822	-5.428	0.000	***
Wordsearch	-1.322	0.383	1,715.750	-3.452	0.001	***
Location:UCL	14.011	1.173	76.382	11.950	0.000	***

Significance. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant	(Intercept)	22.025	4.693
Residual		17.184	4.145

Model Fit Statistics:

Statistic	Value
AIC	10,255.027
BIC	10,309.741
Log Likelihood	-5,117.514
Deviance	10,235.027

6.9 Lane deviations by Scenario and Activity LMM Results

A Generalized Linear Mixed Model (GLMM) was fitted to assess the likelihood of lane deviations (swerve) as a function of activity and scenario type. The model includes activity and scenario type as fixed effects to evaluate how these factors influence the probability of a lane deviation. Location and participant are treated as random effects to account for variability across different testing locations and individual differences in driving behaviour.

Given the binary nature of the outcome (lane deviation or no lane deviation), a binomial distribution with a logit link function was used to model the probability of lane deviations (swerve). This approach helps to analyse how these factors affect lane deviations while controlling for individual and location-based variability.

The results provide insight into how different NDRAs and driving scenarios impact the likelihood of lane deviations, helping to understand driver performance under varying conditions. Full model details and results are available in Technical Appendix 6.9.

Fixed Effects:

Model Formula: Swerve ~ Activity + Scenario Type + (1 | Location/Participant)

Effect	Estimate	Standard Error	z-value	p-value	Significance
(Intercept)	-1.245	0.912	-1.366	0.172	
Drinking water	-0.080	0.394	-0.204	0.838	
Eating popcorn	-0.046	0.395	-0.116	0.907	
Handheld mobile phone	-0.660	0.400	-1.652	0.099	.
Magazine	-0.295	0.391	-0.755	0.450	
No NDRA (Congestion)	-0.218	0.410	-0.532	0.595	
No NDRA (Roadworks)	0.046	0.405	0.114	0.909	
Wordsearch	-0.525	0.395	-1.330	0.183	
Scenario Type Roadworks	0.428	0.229	1.873	0.061	.

Significance. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant:Location	(Intercept)	2.115	1.454
Location	(Intercept)	1.397	1.182

Model Fit Statistics:

Statistic	Value
AIC	733.542
BIC	782.721
Log Likelihood	-355.771
Deviance	550.441

6.10 Looking time in mirrors LMM Results

A linear mixed-effects model (LMM) was used to examine whether mirror looking time differed by activity and Area of Interest. The model included activity and area as fixed effects, while location and participant were treated as random effects to account for variability between locations and individuals. This approach allows for an analysis of how different activities influenced the time participants spent looking in mirrors across various areas.

Fixed Effects:

Model Formula: Looking Time ~ Activity + Area + (1 | Location/Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
(Intercept)	0.170	0.049	372.572	3.441	0.001	***
Drinking water	0.133	0.051	1,288.211	2.607	0.009	**
Eating popcorn	0.025	0.051	1,287.437	0.480	0.631	
Handheld mobile phone	0.063	0.051	1,288.211	1.234	0.217	
Magazine	0.138	0.051	1,287.866	2.706	0.007	**
No NDRA (Congestion)	0.167	0.051	1,287.866	3.280	0.001	***
No NDRA (Roadworks)	0.035	0.051	1,287.866	0.682	0.495	
Wordsearch	0.058	0.051	1,287.866	1.141	0.254	
Right side mirror	0.099	0.025	1,287.099	3.902	0.000	***

Significance. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Random Effects:

Group	Random Effect	Variance	Std. Deviation
Participant:Location	(Intercept)	0.083	0.289
Location	(Intercept)	0.000	0.000
Residual		0.223	0.472

Model Fit Statistics:

Statistic	Value
AIC	2,074.862
BIC	2,137.637
Log Likelihood	-1,025.431
Deviance	2,050.862

6.11 Location and Looking time in mirrors LMM Results

A linear mixed-effects model (LMM) was fitted to examine the effects of Activity, Area, and Location on Looking Time in mirrors only. In this model, NDRA, AOI, and Location are treated as fixed effects to assess their individual and combined influences on the time participants spend looking in mirrors. Participant is included as a random effect to account for variability between individuals. This model allows for the evaluation of how different activities and areas, as well as the testing location, impact looking time while controlling for participant-specific differences.

Fixed Effects:

Model Formula: Looking Time ~ Activity + Area + Location + (1 | Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
(Intercept)	0.152	0.074	146.844	2.063	0.041	*
Drinking water	0.133	0.051	1,288.192	2.607	0.009	**
Eating popcorn	0.025	0.051	1,287.425	0.480	0.631	

Model Formula: Looking Time ~ Activity + Area + Location + (1 | Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
Handheld mobile phone	0.063	0.051	1,288.192	1.234	0.217	
Magazine	0.138	0.051	1,287.845	2.706	0.007	**
No NDRA (Congestion)	0.167	0.051	1,287.845	3.279	0.001	***
No NDRA (Roadworks)	0.035	0.051	1,287.845	0.682	0.496	
Wordsearch	0.058	0.051	1,287.845	1.140	0.254	
Right side mirror	0.099	0.025	1,287.092	3.902	0.000	***
Location:UCL	0.025	0.075	84.959	0.333	0.740	

Significance. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant	(Intercept)	0.084	0.290
Residual		0.223	0.472

Model Fit Statistics:

Statistic	Value
AIC	2,078.092
BIC	2,140.867
Log Likelihood	-1,027.046
Deviance	2,054.092

6.12 Pupil diameter change rate LMM Results

A linear mixed-effects model (LMM) was fitted to examine the pupil diameter change rate based on scenario type and activity. The model included location and participant as random effects to account for individual variability and differences between locations. This model helps evaluate how scenario type and activity influence changes in pupil diameter over time.

Fixed Effects:

Model Formula: Pupil Diameter Change Rate ~ Scenario Type + Activity + (1 | Location/Participant)

Effect	Estimate	Standard Error	df	t value	Pr(>|t|)	Significance
(Intercept)	0.046	0.006	178.831	7.479	0.000	***
Scenario Type Roadworks	-0.012	0.003	3,896.052	-4.322	0.000	***
Drinking water	0.005	0.005	3,867.284	1.019	0.308	
Eating popcorn	-0.006	0.005	3,871.932	-1.265	0.206	
Handheld mobile phone	-0.035	0.005	3,871.271	-7.611	0.000	***
Magazine	-0.007	0.005	3,880.342	-1.436	0.151	
No NDRA (Congestion)	0.047	0.005	3,867.146	9.894	0.000	***
No NDRA (Roadworks)	0.008	0.005	3,872.219	1.475	0.140	
Wordsearch	-0.046	0.005	3,878.279	-9.692	0.000	***

Significance. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant:Location	(Intercept)	0.002	0.046
Location	(Intercept)	0.000	0.000
Residual		0.005	0.072

Model Fit Statistics:

Statistic	Value
AIC	-9,173.502
BIC	-9,098.131
Log Likelihood	4,598.751
Deviance	-9,197.502

6.13 Pupil diameter Interaction LMM Results

A linear mixed-effects model was used to investigate the interaction between activity and location on the pupil diameter change rate. This model includes scenario type as a fixed effect and tests whether the relationship between activity and pupil diameter change rate varies across different locations. Participant is included as a random effect to control for individual differences.

Fixed Effects:

Model Formula: Pupil Diameter Change Rate ~ Scenario Type + Activity * Location + (1 | Participant)

Effect	Estimate	Standard Error	df	t value	Pr(>|t|)	Significance
(Intercept)	0.060	0.011	163.024	5.252	0.000	***
Scenario Type Roadworks	-0.014	0.003	3,886.119	-5.202	0.000	***
Drinking water	-0.019	0.009	3,861.690	-2.133	0.033	*
Eating popcorn	-0.034	0.009	3,872.254	-3.675	0.000	***
Handheld mobile phone	-0.026	0.009	3,868.338	-2.979	0.003	**
Magazine	0.010	0.009	3,875.351	1.114	0.265	
No NDRA (Congestion)	-0.008	0.009	3,867.302	-0.943	0.346	
No NDRA (Roadworks)	0.009	0.009	3,871.012	0.940	0.347	
Wordsearch	-0.036	0.009	3,875.570	-4.076	0.000	***

Model Formula: Pupil Diameter Change Rate ~ Scenario Type + Activity * Location + (1 | Participant)						
Effect	**Estimate**	**Standard Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
Location:UCL	-0.018	0.013	158.569	-1.384	0.168	
Drinking water: LocationUCL	0.033	0.0105	3,860.937	3.196	0.001	***
Eating popcorn: Location:UCL	0.0370	0.011	3,869.492	3.459	0.001	***
ActivityHandheld mobile phone: Location:UCL	-0.012	0.010	3,866.401	-1.204	0.229	
Magazine: Location:UCL	-0.024	0.010	3,874.491	-2.350	0.019	*
No NDRA (Congestion): Location:UCL	0.073	0.010	3,867.009	7.234	0.000	***
No NDRA (Roadworks): Location:UCL	-0.000	0.010	3,867.521	-0.008	0.994	
Wordsearch: Location:UCL	-0.014	0.010	3,873.780	-1.385	0.166	

Significance. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Random Effects:

Group	Random Effect	Variance	Standard Deviation
Participant	(Intercept)	0.002	0.046
Residual		0.005	0.071

Model Fit Statistics:

Statistic	Value
AIC	-9,250.024
BIC	-9,130.686

Statistic	Value
Log Likelihood	4,644.012
Deviance	-9,288.024