International Journal of Human - Computer Studies xxx (xxxx) 103554

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

# Investigating drivers' responses to cyber-attacks while conducting non-driving related tasks in highly automated vehicles

Gayoung Ban, Myounghoon Jeon *

*Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061, United States*

ARTICLE INFO

*Keywords:*
Automated vehicles
Cognitive workload
Cybersecurity
Human factors
Situation
Awareness
Eye gaze behavior

ABSTRACT

As automated vehicles (AVs) advance, understanding human factors in cybersecurity incidents is essential to ensuring driver safety and system resilience. While prior research has explored driver responses to cyber-attacks in partially automated (Level 2–3) vehicles, less is known about how drivers in highly automated vehicles respond. In Level 4 automation, drivers are not required to monitor the roadway continuously but may still need to intervene in unforeseen cyber-attack, making re-engagement dynamics fundamentally different from lower levels of automation. This study examines the impact of non-driving-related task (NDRT) engagement and cyber-attack criticality on situation awareness, visual attention, response time, and workload in Level 4 AVs. To this end, forty-five participants drove in a driving simulator with two types of cyber-attack criticality (non-safety-related, and safety-related as within-subjects) and three non-driving related tasks (NDRTs) engagement levels (no, single and dual as between-subjects). Results indicate that drivers engaged in any level of NDRT (Single or Dual) had significantly reduced situation awareness of road conditions and delayed response time and gaze reallocation to critical information after a cyber-attack, particularly in Dual NDRT conditions. Additionally, safety-related cyber-attacks induced greater cognitive workload, suggesting that drivers exert more mental effort when responding to high-risk threats. These findings highlight the unique re-engagement challenges in Level 4 AVs, where drivers must transition from passive engagement in NDRTs to active situation awareness during cybersecurity incidents. The results emphasize the need for human-centered AV cybersecurity systems that optimize alert delivery, minimize cognitive overload, and facilitate rapid driver response to emerging threats in highly automated driving environments.

## 1. Introduction

Automated vehicles (AVs), also known as autonomous or self-driving vehicles (Greenblatt and Shaheen, 2015) aim to reduce human involvement in driving by using advanced sensors and algorithms for decision-making. Core components like motion control, environmental sensing, and path planning (Pendleton et al., 2017) enable AVs to operate independently, improving traffic safety and efficiency while playing a key role in intelligent transportation systems (ITSs). To standardize the classification of vehicle automation, the Society of Automobile Engineers (SAE) International developed the Levels of Driving Automation framework. This framework identifies six levels of automation, ranging from Level 0, representing no automation, to Level 5, which denotes full automation under all conditions (SAE, 2016). Sheehan et al. (2017) emphasize that as automation increases, the responsibility for driving progressively shifts from the human driver to the automated system.

Level 4 vehicles, classified as High Automation, can operate without human intervention in most circumstances. Unlike Level 3 vehicles, which require driver intervention in certain situations, Level 4 vehicles can perform all driving tasks independently, with human input needed only in particularly complex or unanticipated environments. Companies like Tesla, GM, Ford, Uber, and Google are advancing autonomous systems, driving the commercialization of ride-hailing and robotaxi services (Yao et al., 2020; Zhou and Xu, 2023). Some operators are already testing autonomous services, with predictions that taxi and ride-hailing will be predominantly autonomous within the next decade (Golbabaei et al., 2021). Forecasts indicate significant market penetration for Level 4 AVs, reaching 15 % by 2030 and 20 % by 2050 (Nieuwenhuijsen et al., 2018), with full adoption expected after 2070. Bansal and Kockelman (2017) estimate penetration rates between 24.8 % and 87.2 % by 2045, depending on costs and consumer adoption.

* Corresponding author at: Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061, United States.
*E-mail addresses:* gayoungban@vt.edu (G. Ban), myounghoonjeon@vt.edu (M. Jeon).

https://doi.org/10.1016/j.ijhcs.2025.103554
Received 5 October 2024; Received in revised form 19 May 2025; Accepted 20 May 2025
1071-5819/

However, as automation increases, so does the reliance on complex computational systems, leaving AVs vulnerable to cybersecurity risks (As cited in Siddiqui et al., 2023). Cyber-attacks on such highly automated systems (Level 4 vehicles) pose serious threats by manipulating safety-critical information or disrupting vehicle operations, creating potentially hazardous situations (Rofail et al., 2019). Such attacks often target the In-Vehicle Information System (IVIS) by modifying or falsifying critical information, such as false congestion alerts, tampered collision warnings, or bogus speed notifications, misleading drivers and affecting their situation awareness (Parkinson et al., 2017). These attacks are facilitated by vulnerabilities in-vehicle sensors and communication protocols, making them prime targets for spoofing, jamming, and Denial of Service (DoS) attacks. For instance, cameras and LiDAR sensors are particularly susceptible to jamming attacks, which block communication signals and disrupt the vehicle's perception of its environment (Petit et al., 2015; Thing and Wu, 2016; Meryem and Mazri, 2019).

In Level 4 vehicles, if a cyber-attack successfully manipulates communication channels or disrupts critical systems, the vehicle becomes entirely vulnerable, with no immediate human oversight to intervene. This contrasts with Level 3 vehicles, where human drivers are required to remain available and can regain control during emergencies. Moreover, such perception-based cyber-attacks introduce serious cognitive and perceptual challenges. These attacks directly affect the driver's ability to interpret the situation and act accordingly. For example, it can create confusion, reduce situation awareness (SA) of the road environment, and increase mental workload, leading to delayed responses and decision-making errors, particularly when they occur without sufficient warning. Driver cognitive engagement is another critical challenge in Level 4 vehicles due to frequent involvement in non-driving-related tasks (NDRTs). While NDRTs are allowed in highly automated driving, they can result in cognitive detachment from the driving task, impairing driver readiness to respond to unexpected cyber-attacks in complex environments (Andrade and Yoo, 2019; Noy, 2018). Perception-based cyber-attacks target cognitive processes directly affecting the driver's ability to interpret the situation and act accordingly. As a result, understanding how drivers perceive and respond to these attacks is crucial for ensuring safety in highly automated vehicles. Therefore, understanding how cyber-attacks affect driver response behavior in Level 4 vehicles is essential for ensuring safety in fully automated environments.

## 2. Related works

### 2.1. Cybersecurity threats in automated vehicles

Prior research on cybersecurity in automated vehicles (AVs) has extensively explored technical detection algorithms and prevention strategies (Ahmad et al., 2019; Eiza and Ni, 2017). While these studies provide critical insights into system vulnerabilities, they largely overlook the human factors involved in responding to cyber-attacks once they occur. This gap is particularly critical in Level 4 automation, where drivers are permitted to disengage from active vehicle control, making re-engagement and response to cybersecurity threats more complex.

### 2.2. Driving performance and behavioral adaptations under cyber-attacks

Empirical studies using driving simulators have begun to explore driving performance during cyber-attacks in AVs. Dong et al. (2024) focused on lane-changing behavior under cyber-attacks in fully automated connected vehicles (CAVs) that follow predefined car-following and lane-change algorithms without any human interventions. Their study modeled the impact of remote-access cyber-attacks on traffic flow and safety, revealing that cyber-attacked vehicles exhibited more erratic lane changes, leading to increased congestion and accident risks. A related study by F. Zhang et al. (2023) explored how driving style affects driving behaviors to unexpected cyber-attacks using a SAE Level 2 driving simulator. The study measured braking reaction time, post-event acceleration, and time to first reaction under different attack scenarios. Their results revealed that drivers with higher sensation-seeking tendencies exhibited faster responses and more controlled acceleration after cyber-attacks, contrary to prior assumptions that risk-prone drivers would respond less cautiously. This finding suggests that individual differences in driving style and risk perception may influence cyber-attack response effectiveness, highlighting the need for personalized cybersecurity intervention strategies in AVs. He et al. (2024) investigated the impact of distraction-based cyber-attacks on driving performance, comparing their effects to alcohol-induced impairment, using a simulator configured with a Level 2 driver-assist system. Their findings showed that cyber-attacks imposing high cognitive demand led to significant delays in braking reaction times and increased lane deviation errors, indicating that the mental burden of processing cybersecurity threats can severely impair driver performance. These results emphasize the importance of managing cognitive load in automated driving, as excessive mental demand may hinder timely and effective responses to cyber-attacks. Aliebrahimi and Miller (2023) examined how cybersecurity knowledge and situation awareness influenced driving performance under cyber-attacks. In an SAE Level 3 conditional-automation simulator—supplemented by a post-experiment survey, they found that drivers with greater cybersecurity awareness exhibited improved situation awareness and shorter takeover times when transitioning from automated to manual control. However, their study also revealed that some drivers, despite recognizing cyber-attacks, failed to take corrective action, highlighting a critical gap between awareness and intervention.

### 2.3. Visual attention allocation in cyber-attack detection

Visual attention allocation plays a crucial role in how drivers detect and respond to cybersecurity threats in automated vehicles (AVs). Understanding how gaze behavior adapts to system failures and cyber-attacks is essential for designing effective cybersecurity interventions.

Louw et al. (2019) investigated drivers' fixation patterns and gaze distribution as they were repeatedly exposed to "silent" automation failures, where the system ceased providing feedback without issuing a takeover request—in a SAE Level 3 high-fidelity driving simulator. At first exposure, drivers' gaze was predominantly drawn to the instrument-cluster area where the failure occurred, resulting in prolonged fixations and delayed reallocation of attention to the forward roadway. However, with successive failure exposures, fixation durations on the cluster cue steadily decreased, and gaze distribution became more evenly spread across critical road regions, indicative of a learning or adaptation effect. Over time, drivers increasingly directed their attention to areas most relevant for safe vehicle operation, enhancing both situation awareness and cognitive flexibility. Although this study did not involve malicious cyber-attacks, these results suggest that structured, repeated exposure to cybersecurity-related anomalies in AVs could similarly train drivers to detect and respond more effectively to cyber threats. Payre et al. (2023) further demonstrated how cybersecurity failures, such as ransomware attacks and silent system malfunctions, impact driver distraction in a conditional automation (SAE Level 3) setting. Their findings revealed that drivers exposed to explicit failures, such as ransomware messages on the in-vehicle screen, exhibited prolonged visual fixations on the alert, with some drivers looking at the ransomware for over 12 s. This excessive visual engagement diverted attention from the road and increased safety risks, highlighting the potential dangers of poorly designed cybersecurity alerts that demand excessive driver focus.

The effectiveness of cybersecurity training in improving visual attention strategies has also been explored. Wang et al. (2024) investi-

gated the impact of training and warning systems on driver responses to cyber-attacks in the SAE Level 3 (conditional-automation), dividing participants into four conditions: no intervention, training-only, warning-only, and combined training and warning. Their findings revealed that training significantly improved driver preparedness, leading to more frequent visual scanning and timely braking responses. However, real-time warnings alone were often insufficient, as some drivers either ignored or misinterpreted alerts, emphasizing the need for pre-exposure cybersecurity training rather than relying solely on in-the-moment warnings.

F. Zhang et al. (2023) examined drivers' visual attention allocation and response behaviors under different cyber-attack scenarios, including false emergency sirens, manipulated dashboard warning messages, and unintended lane changes using the SAE Level 2 (partial automation). Their findings demonstrated that attack visibility played a crucial role in determining driver responses. Participants who received explicit auditory and visual warnings directed their gaze more frequently toward the IVIS and road, enabling faster braking reaction times and timely steering adjustments. Conversely, drivers subjected to silent system manipulations exhibited prolonged fixations on secondary tasks or delayed gaze shifts, leading to slower or absent intervention.

### 2.4. Trust and cybersecurity in automated vehicles

Another key concern in AV cybersecurity is its impact on trust in automation. Lim et al. (2024) demonstrated that cyber-attacks significantly degrade driver trust in an SAE Level 3 (conditional automation) simulator, with trust remaining impaired even after a successful, attack-free driving session. Similarly, Marcinkiewicz and Morgan (2023) conducted an experiment using the video-based Level 3 automation to assess how cyber readiness and response strategies of an autonomous vehicle (AV) company influence driver trust and blame assignment following a cyber-attack. The study measured trust ratings in both the AV company and the AV itself before and after the cyber-attack. The results showed a significant decline in trust immediately following the attack, even before participants were informed of the company's response. However, after learning about how the company handled the incident, trust levels varied depending on the company's level of cyber readiness and the nature of their response.

### 2.5. Research gaps and study contributions

While these studies provide valuable insights into drivers' behavioral and cognitive adaptations under cybersecurity threats, several key research gaps remain. Existing studies do not adequately explore the role of non-driving-related task (NDRT) engagement in shaping driver response to cyber-attacks. In Level 4 AVs, where drivers are permitted to engage in NDRTs, it remains unclear how different levels of task engagement (No NDRT, Single NDRT, Dual NDRT) affect situation awareness, gaze behavior, response time, and workload during cyber-attacks. Additionally, while some studies classify cyber-attacks based on their impact on AV functionality, little empirical work has examined how cyber-attack criticality (safety-related vs. non-safety-related) affects driver responses.

### 2.6. Research questions and hypotheses

This study addresses these gaps by investigating how NDRT engagement levels and cyber-attack criticality interact to shape driver situation awareness, visual attention allocation, response time, and cognitive workload in Level 4 AVs. We also derive practical design implications—grounded in a post-experiment driver preference survey and in prior simulator studies of AV cyber-attacks (Dong et al., 2024; Zhang et al., 2023; He et al., 2024; Payre et al., 2023)—to inform the development of more effective in-vehicle cybersecurity alert systems.

The present research questions and hypotheses are

**H1**. In a Level 4 automated vehicle experiencing a cyber-attack, how do No NDRT, Single NDRT, and Dual NDRT engagement levels affect drivers' situation awareness, gaze behavior, response time, and workload?

**H1a**. Drivers engaged in Single NDRT and Dual NDRT will exhibit lower situation awareness scores compared to drivers with No NDRT

**H2b**. Drivers in Dual NDRT conditions will exhibit longer glance durations toward the IVIS or road environment after a cyber-attack occurs, compared to drivers in Single NDRT or No NDRT conditions

**H1c**. Drivers in Single NDRT and Dual NDRT engagement conditions will have longer response time during cyber-attacks

**H1d**. Drivers experiencing Single NDRT and Dual NDRT engagement will report higher cognitive workload after a cyber-attack

**H2**. How do cyber-attack criticality levels affect drivers' situation awareness, gaze behavior, response time, and workload?

**H2a**. During safety-related cyber-attacks, drivers will exhibit lower situation awareness scores compared to non-safety-related cyber-attacks

**H2b**. During safety-related cyber-attacks, drivers will gaze longer toward the IVIS screen and road compared to non-safety-related cyber-attacks

**H2c**. During safety-related cyber-attacks, drivers will exhibit faster response time compared to non-safety-related cyber-attacks

**H2d**. Drivers will exhibit higher cognitive workload after experiencing safety-related cyber-attacks compared to non-safety-related cyber-attacks

## 3. METHOD

### 3.1. Experimental design and variables

The study employed a 3 × 2 mixed design, with two independent variables: NDRT engagement level (No NDRT, Single NDRT, and Dual NDRT) and cyber-attack criticality level (non-safety-related and safety-related scenarios). The NDRT engagement level was a between-subjects variable, while the cyber-attack criticality level was a within-subjects variable. Participants completed an experimental task that included both an NDRT phase and a subsequent response task to a cyber-attack phase.

As for the dependent variables, situation awareness, eye gaze behavior, response time, and perceived workload measures were employed (Table 1). Situation awareness (SA) was measured by the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 2000). SAGAT consisted of two queries for each level of SA: Level 1 (perception), Level 2 (comprehension), and Level 3 (projection), and each query was made based on the specific events presented on the road (see surve B). The average score for each SA level was calculated for each participant. Moreover, the total average score of SA was calculated from the average score of the three SA levels. Eye gaze behavior was examined using the measure of total glance duration. Also, eye gaze behavior during the response task was examined by calculating the percentage of time spent in each AOI. This approach allowed us to capture visual attention allocation across different conditions and compensate for variability in response behaviors, as participants' reaction times and strategies naturally varied depending on the type and criticality of the cyber-attack. Response time was determined as the time duration between the cyber-attack issuance and the onset of the driver's maneuver (the time when the participant started hitting the pedal). The workload measure employed was the NASA-TLX scores (Hart, 1988). For NDRT

**Table 1**
Dependent measures.

| Dependent Measure | Definition | Assessment Method |
|---|---|---|
| Situation Awareness (SA) | The driver's understanding of the current and future system state. | Six SAGAT queries (2 per level: perception, comprehension, projection). |
| Eye Gaze Behavior | Visual attention allocation among areas of interest (AOI). Four AOIs were considered: road, IVIS screen, dashboard, and NDRT interface (Fig. 1, Left). | Total glance duration. It is defined as the sum of the lengths of the individual glances to an area of interest (AOI). The length of a single glance is the time from the start of the saccade leading into the AOI to the end of the last fixation on the AOI. |
| Response Time | Latency from alert onset to driver action. | The time duration between the cyber-attack issuance and the onset of the driver's maneuver (the time when the participant started hitting the pedal). This measure was determined using the log data from the driving simulation software program. |
| Workload | Subjective mental workload during cyber-attack response. | NASA-TLX questionnaire. |

performance, the results of the typing test were used. The number of errors in typing was calculated and analyzed as a performance measure.

### 3.2. Participants

Forty-five participants (27 males and 18 females) were recruited and divided into three groups, 15 each as the No NDRT group, Single NDRT group, and Dual NDRT group, according to the NDRT engagement levels. Their mean age and driving experience were 25.2 years ($SD$ = 3.64, min = 20, max = 35) and 5.63 years ($SD$ = 2.53, min = 1, max = 14), respectively. Participants reported their familiarity with Level 2 partial-automation features—adaptive cruise control (ACC) and lane-keeping assist (LKA) —at a mean of 3.20 ($SD$ = 0.98, min = 2, max = 4) (on a 1 (no experience) to 5 (high experience) scale. All participants had normal or corrected-to-normal vision and held a valid driver's license for at least 6 months. All participants were paid USD 10 for their participation. The study received ethical approval from the Institutional Review Board (Virginia Tech IRB-23–1102).

### 3.3. Apparatus

The Driving Safety Research Institute's (DSRI) miniSim fixed-base driving simulator, consisting of three 42-inch widescreen monitors, a dashboard, and a PC, was used in this study (Fig. 1). The simulated vehicle is equipped with Level 4 (*SAE*, 2016) Automation, capable of autonomous operations such as lane-keeping and adaptive cruise control. A 16-inch laptop was employed to simulate an NDRT interface, and a 10.9-inch iPad was employed as the central console to represent an in-vehicle infotainment system (IVIS) screen. The IVIS screen displayed a static navigation map view on the right side and a dynamic vehicle status view, such as the mode of control (i.e., manual or automated), and alerted drivers by presenting visual cyber-attack messages with an auditory alarm (a beeping sound) (Fig. 1). Each participant's driving data were recorded with a sampling rate of 60 Hz. Each participant's eye movements were recorded using a Tobii Pro Glasses 2 eye-tracker with a sampling rate of 50 Hz. Eye-tracking dataset was coded and analyzed using the Tobii Pro Lab® Software version 1.171.

### 3.4. Procedure

Prior to the experimental trials, participants underwent an introduction and training session. During the introduction, participants were informed that the study simulated Level 4 AV, which allows drivers to engage in non-driving-related activities while the vehicle handles all driving tasks. They were encouraged to interact with the environment as they would in a real-world Level 4 vehicle, but were advised not to fall asleep. The training session was designed to familiarize participants with the driving simulator and the experimental tasks. The experimental task consisted of two successive sub-tasks: an NDRT followed by a response task to cyber-attacks (Fig. 2).

The NDRTs used in this study varied according to the level of NDRT engagement. These tasks were selected to reflect real-world activities that drivers might engage in while operating a vehicle. Ko and Ji (2018) Argue that natural tasks can be more effective than controlled artificial tasks in understanding driver behavior in automated driving scenarios. Therefore, the NDRTs employed in our study aimed to replicate real-life tasks that drivers might encounter.

In the No NDRT group, participants were not engaged in any NDRT. Participants were instructed to engage freely with their surroundings while remaining seated and attentive. Although they were not required


**Fig. 1.** Experimental setup with four AOIs (left). The four AOIs are on the left: (a) road, (b) dashboard, (c) IVIS screen, and (d) NDRT interface.

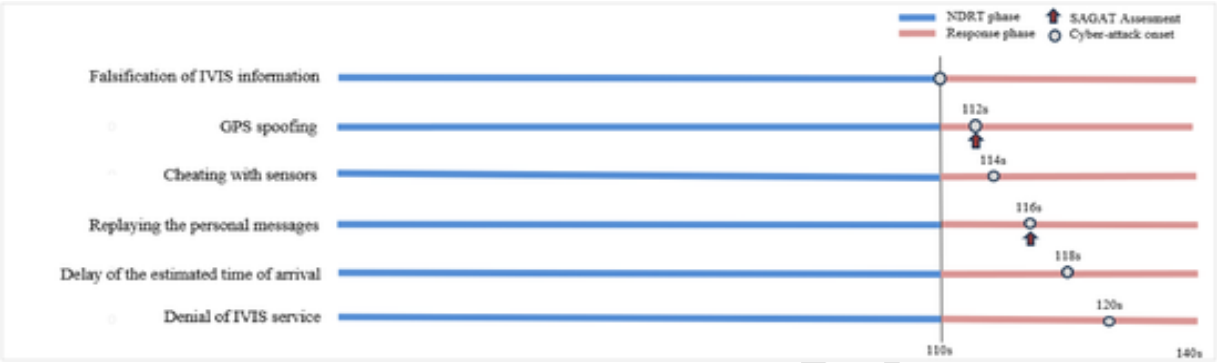



**Fig. 2.** Experimental procedure.

to focus on the driving environment continuously, they were reminded not to fall asleep and to remain aware of potential alerts. In the Single NDRT group, participants were instructed to engage in a typing task (using an Internet-based typing practice application) on a laptop placed on their knees (Fig. 1). This task involved visual and manual demands, requiring participants to look down at the laptop, limiting their visual input from the surrounding environment. Detailed instructions on how to perform the task were provided during the training session. They were informed that they could focus primarily on the typing task but should still remain passively aware of the driving environment. The Dual NDRT group performed the typing task on a laptop placed on their knee while simultaneously engaging in a spoken word-chain game. In this game, participants verbally generated words aloud starting with the last letter of the previous word provided by the experimenter (e.g., "apple" → "elephant" → "tiger"), simulating real-world multitasking such as talking to passengers while composing emails. During the training session, they were guided on how to manage the dual tasks effectively. In this condition, participants were not required to monitor the driving environment actively, reflecting real-world scenarios where occupants in highly automated vehicles are fully engaged in NDRTs. Immediately after performing the NDRT, the response task to cyber-attacks began, with a cyber-attack message appearing on the IVIS screen, accompanied by a simultaneous 2-second auditory alarm (a beeping sound). Visual messages to IVIS screen were used in this study to counteract the invisibility of cyber-attacks, which are often overlooked due to their lack of direct sensory impact (Pyke et al., 2023). Unlike physical threats that provide tangible cues, cyber threats occur silently in the background, making them less likely to be detected and anticipated, especially when drivers are engaged in NDRTs. The experiment consisted of six trials, each presenting a unique cyber-attack message on the IVIS screen. In each trial, participants responded by pressing the pedal—our primary input for capturing reaction time. Although different messages could map to braking, takeover confirmation, alert acknowledgment, or IVIS reboot, we uniformly measured the latency from message onset to pedal press across all six attack scenarios. While the pedal had multiple functions, the primary focus was on measuring real-time reaction time to diverse cyber-attack scenarios. Steering wheel interactions were not considered as a response option because participants' hands were frequently occupied with NDRTs, which could have compromised the consistency and accuracy of steering-based responses. The time budget for the response task to cyber-attacks was set at 30 s (Zhang et al., 2019).

At the beginning of each experimental trial, the simulated ego vehicle was self-driving in the second lane of a four-lane highway at a fixed speed of 55 mph. Participants began each trial with 110 s of NDRT engagement, followed by a 30-second response window (B. Zhang et al., 2019) after the onset of each cyber-attack. The cyber-attacks occurred at predetermined times—110 s, 112 s, 114 s, 116 s, 118 s, and 120 s (Fig. 2). These cyber-attacks did not occur all at once but were distributed across trials. To assess participants' situation awareness (SA), SAGAT assessments were conducted during the second and fourth cyber-attacks at 112 s and 116 s, respectively. The driving simulation was paused at these moments without prior warning to the participants. Following each experimental trial, participants completed the NASA Task Load Index questionnaire to assess their workload. The driving simulation was then reinitialized, and the participant's vehicle was regenerated in the second lane for the next trial. Upon completing the experiment, participants were asked to answer short survey questions about their expectations regarding cyber-attack alert display types on the IVIS during the cyber-attacks.

### 3.5. Driving and cyber-attack scenarios

Driving scenarios took place in a 3D virtual replica of the rural area of the USA, which included 10 miles of suburban roads. Cyber-attacks employed in the current study represent real-world threats to current vehicle technology (Wang et al., 2008). The scenarios presented in this study were designed as perception-based manipulations rather than physical changes to vehicle dynamics. Although the actual vehicle behavior remained unaffected, these manipulations altered the information displayed on the in-vehicle touchscreen, leading drivers to respond based on their criticality level. Non-safety-related scenarios focus on spying and manipulating drivers' private information stored in the infotainment system of vehicles. These scenarios involve the misrepresentation of secondary information that does not directly impact the vehicle's control or safety-critical functions but can divert the driver's attention, increase cognitive load, and reduce situation awareness. In such cases, drivers are more likely to engage in verification behaviors, such as repeatedly glancing at the touchscreen to confirm the accuracy of the displayed information. On the other hand, safety-related scenarios differ from non-safety-related ones by presenting critical information that appears directly related to vehicle control or driving safety, prompting drivers to respond with immediate physical interventions. Although these scenarios were based on manipulated information rather than actual changes to vehicle dynamics, the perceived threat level led participants to behave as if the vehicle was malfunctioning. This resulted in corrective actions such as braking, rerouting, or frequent monitoring of the dashboard. Each category has three different cyber-attack scenarios, yielding a total of six different cyber-attack scenarios. To minimize predictability and reduce the potential influence of training effects on behavior, cyber-attack alerts were introduced at varying times (Fig. 2) within the final 30 s of the driving scenario. The messages were kept consistent at around 25 words to ensure uniformity in cognitive demand. Table 2 outlines the specifics of each scenario, and Fig. 3 showcases the cyber-attack alerts displayed on the IVIS screen, which participants respond to during the experiment.

**Table 2**
Cyber-attacks scenarios.

| Cyber-attack criticality levels | Scenario Name | Description |
|---|---|---|
| **Non-safety-related scenarios** | Falsification of IVIS information | A hacker alters the IVIS system information, causing it to display a different driver's name and phone model during a connection. |
| | Replaying the personal messages | A hacker infiltrates a driver's text messaging app, accessing previous messages, and systematically resending them to the intended recipient, prompting the issuance of confirmation alert messages for message delivery. |
| | Delay of the estimated time of arrival | A hacker floods the automated vehicle's communication network, causing navigation data disruption. This leads to a significant delay, alerting the driver with a message of a 45-minute delay to their destination. |
| **Safety-related scenarios** | Cheating with sensors | A hacker obtains speedometer sensor data and changes the perceived speed to 60 mph. Consequently, the vehicle issues an alert message to the driver to reduce speed. |
| | GPS spoofing | A hacker gains unauthorized access to the vehicle's GPS data and alters the destination information. Consequently, drivers receive a destination confirmation alert message with an altered address. |
| | Denial of IVIS service | A hacker freezes the IVIS screen, disrupting vital vehicle information such as speed, direction, and fuel state, leading to an alert message. |

### 3.6. Statistical analysis

A mixed Analysis of Variance (ANOVA) was conducted to test the effects of the independent variables and their interaction on each dependent measure, except for NDRT performance, which was analyzed by an independent samples *t*-test. In cases where the main effect of different NDRT engagement levels was found to be statistically significant, post-hoc analysis with Tukey's Honest Significant Difference (HSD) was conducted to control for the family-wise error rate across multiple comparisons. No missing data were recorded in this study, as all participants completed the assigned tasks and provided full datasets across all measured variables. As a result, no data imputation or exclusion was necessary during the analysis for each ANOVA; the homogeneity of data was tested using Levene's test. In cases where homogeneity was violated, non-parametric tests were used; otherwise. A Welch Analysis of Variance was used to compare the effect of NDRT engagement levels on situation awareness, eye gaze behavior, and NASA-TLX of temporal demand and performance. Post-hoc pairwise comparisons of the independent samples were conducted using the Games-Howell Tests. All statistical tests were conducted at an alpha level of 0.05 using SPSS 26 (IBM Corp., Armonk, USA).

## 4. Results

### 4.1. Situation awareness (SAGAT score)

A series of Welch's ANOVAs were conducted to determine the effect of NDRT engagement levels (No NDRT, Single NDRT, Dual NDRT) on each level of SA scores - Level 1 (perception), Level 2 (comprehension), and Level 3 (projection), as well as in the total average SA score across different levels of NDRT engagement. A significant main effect was observed in each level of SA scores; Level 1, $F(2, 27.50) = 16.78$, $p < .001$ $\omega^2 = 0.41$, Level 2, $F(2, 27.50) = 16.61$, $p < .001$, $\omega^2 = 0.40$, Level 3, $F(2, 27.92) = 16.78$, $p < .001$, $\omega^2 = 0.41$ and, total SA, $F(2, 27.93) = 16.56$, $p < .001$, $\omega^2 = 0.40$ (Table 3). The Games-Howell post-hoc tests revealed that both Single and Dual NDRT groups resulted in a significantly lower SA than the No NDRT group (Table 3). No significant main effect was observed regarding cyber-attack criticality levels. There was no two-way interaction between NDRT engagement levels and cyber-attack criticality levels.

### 4.2. Eye gaze behavior

A series of Welch's ANOVAs were conducted to determine the effect of NDRT engagement levels (No NDRT, Single NDRT, Dual NDRT) on the total glance duration in each AOI: Road, IVIS Screen, Dashboard, and NDRT Interface. Significant main effects were found for all AOIs, indicating that NDRT engagement levels significantly influenced visual attention allocation; road, $F(2, 154.17) = 17.84$, $p < .001$, $\omega^2 = 0.42$, IVIS screen, $F(2, 133.41) = 1118.94$, $p < .001$, $\omega^2 = 0.98$, dashboard, $F(2, 171.47) = 87.26$, $p < .001$ $\omega^2 = 0.48$, and NDRT interface, $F(2, 390.75) = 22.34$, $p < .001$, $\omega^2 = 0.51$ (Fig. 4a). For the response task to the cyber-attack phase, a series of Welch's ANOVAs were conducted to determine the effect of NDRT engagement levels on the percentage of time participants spent in each AOI. Significant main effects were found for all AOIs; road, $F(2,55.45) = 16.18$, $p < .001$, $\omega^2 = 0.21$, IVIS screen, $F(2, 38.70) = 187.05$, $p < .001$, $\omega^2 = 0.82$, dashboard, $F(2, 43.25) = 44.23$, $p < .001$, $\omega^2 = 0.49$, and NDRT interface, $F(2, 150.52) = 284.26$, $p < .001$, $\omega^2 = 0.79$ (Fig. 4b).


**Fig. 3.** Cyber-attacks alerts are displayed on the IVIS screen.

**Table 3**
The main effect of NDRT engagement levels. Post-hoc test results are shown within each row. Shared superscript letters ($^{a,b,c}$) indicate no significant difference from one another ($\alpha = 0.05$).

| | | No NDRT Mean (SD) | Single NDRT Mean (SD) | Dual NDRT Mean (SD) | Test Statistics | p-value |
|---|---|---|---|---|---|---|
| Situation awareness | Level 1 | 88.33 (26.50) | 32.50 (35.29)$^{a}$ | 35 (33.81)$^{a}$ | $F(2, 27.50) = 16.78$ | $P < .001$ |
| | Level 2 | 83.33 (30.86) | 28.33 (33.89)$^{a}$ | 28.33 (31.15)$^{a}$ | $F(2, 27.50) = 16.61$ | $P < .001$ |
| | Level 3 | 81.67 (30.57) | 25 (28.35)$^{a}$ | 26.67 (32)$^{a}$ | $F(2, 27.92) = 16.78$ | $P < .001$ |
| | Total | 84.44 (28.84) | 28.61 (32.15)$^{a}$ | 30 (31.62)$^{a}$ | $F(2, 27.93) = 16.56$ | $P < .001$ |
| Response time | | 3.79 (1.74) | 4.57 (1.52)$^{a}$ | 4.62 (1.85)$^{a}$ | $F(2, 267) = 6.72$ | $P = .014$ |
| Perceived workload | Overall demand | 32.17 (19.84) | 41.80 (19.55)$^{a}$ | 41.35 (23.95)$^{a}$ | $F(2, 267) = 5.91$ | $P = .003$ |
| | Mental demand | 29.50 (20.31)$^{a}$ | 41.44 (25.73)$^{b}$ | 35.33 (27.25)$^{ab}$ | $F(2267) = 5.30$ | $P = .005$ |
| | Temporal demand | 36.67 (23.99)$^{a}$ | 44.83 (25.32)$^{ab}$ | 47.11 (29.52)$^{b}$ | $F(2176.76) = 4.11$ | $P = .017$ |
| | Performance | 31.17 (22.88)$^{a}$ | 39.44 (24.83)$^{ab}$ | 40.28 (28.48)$^{b}$ | $F(2, 267) = 3.51$ | $P = .031$ |

Note. Only statistically significant effects ($p < .05$) are presented.

No significant main effect was observed regarding cyber-attack criticality levels. There was no two-way interaction effect between NDRT engagement levels and cyber-attack criticality levels.

#### 4.2.1. Road

For the NDRT phase, the Games-Howell post-hoc tests revealed that Single NDRT and Dual NDRT groups resulted in a significantly longer glance duration to the road than the No NDRT group. Moreover, the Dual NDRT group showed a significantly longer glance duration to the road than the Single NDRT group (Fig. 4a).

For the response task to the cyber-attacks phase, the Games-Howell post-hoc tests revealed that the Single NDRT group exhibited significantly greater road fixation compared to the No NDRT group, whereas the Dual NDRT group did not show a statistically significant difference from the No NDRT group. Notably, the Dual NDRT group exhibited significantly lower road fixation compared to the Single NDRT group (Fig. 4b).

#### 4.2.2. IVIS screen

For the NDRT phase, the Games-Howell post-hoc tests revealed that the No NDRT group resulted in a significantly longer glance duration to the IVIS screen than Single NDRT and Dual NDRT groups (Fig. 4a).

For the response task to the cyber-attacks phase, the Games-Howell post-hoc tests revealed that the Single NDRT group spent significantly more time looking at the IVIS screen compared to the No NDRT group. However, the Dual NDRT group spent significantly less time looking at the IVIS screen than both the No NDRT and Single NDRT groups (Fig. 4b).

#### 4.2.3. Dashboard

For NDRT phase, the Games-Howell post-hoc tests revealed that No NDRT group resulted in a significantly longer glance duration to IVIS screen than Single NDRT and Dual NDRT groups (Fig. 4a). For response task to cyber-attacks phase, the Games-Howell post-hoc tests revealed that time spent on the dashboard significantly decreased in NDRT-engaged participants. Compared to the No NDRT group, Single NDRT group spent significantly less time checking dashboard information, and Dual NDRT group spent even less time (Fig. 4b).

#### 4.2.4. NDRT interface

For NDRT phase, the Games-Howell post-hoc tests revealed that Single NDRT group and Dual NDRT group spent significantly more time on the NDRT interface compared to No NDRT group. Additionally, the Single NDRT group spent more time on the NDRT interface than the Dual NDRT group (Fig. 4a).

For response task to cyber-attacks phase, the Games-Howell post-hoc tests revealed that Single NDRT group and Dual NDRT group spent significantly more time on the NDRT interface compared to No NDRT group. However, in this case, Dual NDRT group spent more time on the NDRT interface than the Single NDRT group (Fig. 4b).

### 4.3. Response time

A series of ANOVA tests were conducted to determine the effect of NDRT engagement levels (No NDRT, Single NDRT, Dual NDRT) on response time to cyber-attacks. A significant main effect was observed in response time, $F(2, 267) = 6.72$, $p < .05$, $\eta^2 = 0.24$ (Table 3). Tukey's Honest Significant Difference test revealed that Single NDRT and Dual NDRT groups resulted in a significantly longer response time than No NDRT group. For response time measures, no significant main effect was observed regarding cyber-attack criticality levels. There was no two-way interaction effect between NDRT engagement levels and cyber-attack criticality levels.

### 4.4. Perceived workload (NASA-TLX score)

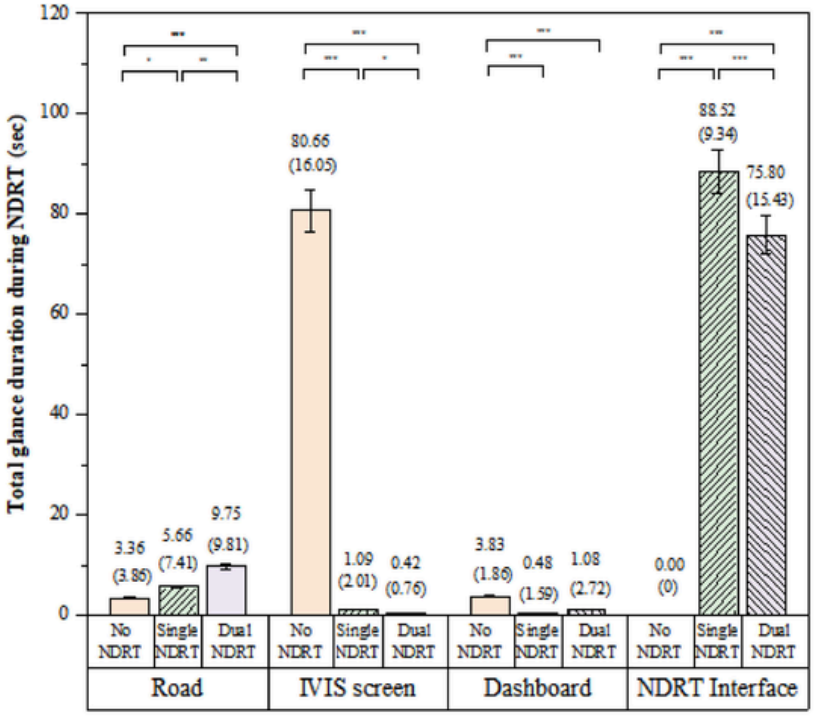
A series of ANOVA tests and Welch's ANOVA test were conducted to determine the effect of NDRT engagement levels (No NDRT, Single NDRT, Dual NDRT) on perceived workload. A significant main effect was observed across different NDRT engagement levels, and a significant interaction effect between NDRT engagement levels and cyber-attack criticality levels was observed. No significant main effect was observed in cyber-attack criticality levels. NDRT engagement levels affected overall workload score, $F(2, 267) = 5.91$, $p < .005$, $\eta^2 = 0.23$, mental demand, $F(2267) = 5.30$, $p < .005$, $\eta^2 = 0.22$, temporal demand, $F(2, 176.76) = 4.11$, $p < .05$, $\omega 2 = 0.29$, and performance, $F(2, 267) = 3.51$, $p < .05$, $\eta^2 = 0.14$. The post-hoc multiple pairwise comparisons indicated that the Dual NDRT group resulted in higher temporal demand, performance, and overall workload scores than the No NDRT group (Table 3). It also revealed that Single NDRT resulted in higher mental demand and overall workload scores than the No NDRT group (Table 3). Cyber-attack criticality levels significantly affected the scores of effort, $F(1, 132) = 4.10$, $p = .044$, $\eta^2 = 0.16$. Safety-related scenarios showed a higher effort than non-safety-related scenarios (Table 4). A significant two-way interaction was found for effort, $F(2, 132) = 3.62$, $p < .05$, $\eta^2 = 0.14$ (Table 4). However, the post-hoc multiple pairwise comparisons revealed no significant differences.

### 4.5. NDRT performance

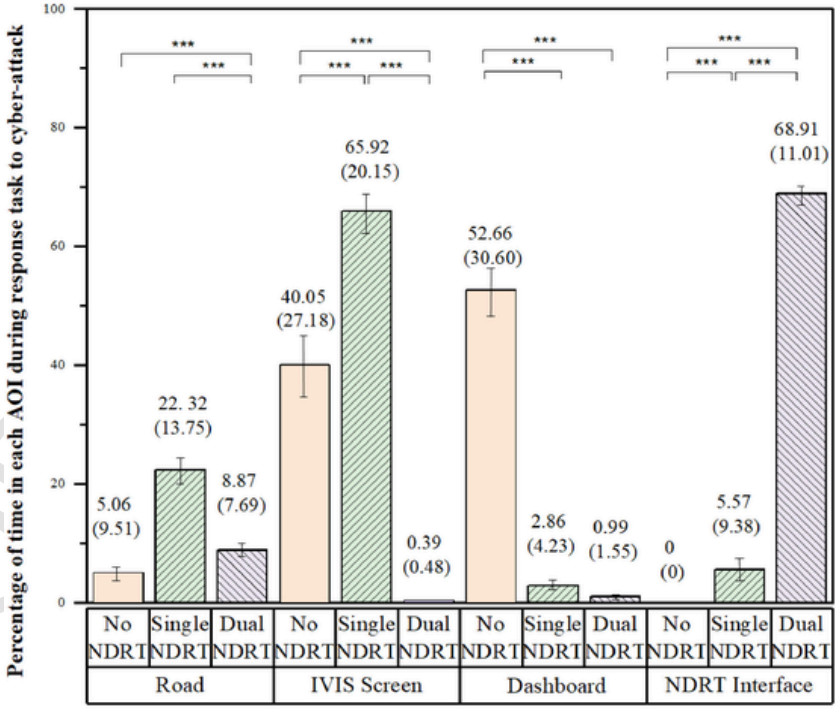
No significant main effect or interaction effect was observed regarding NDRT engagement levels and cyber-attack criticality levels.

### 4.6. User preferences regarding alert systems

A post-experiment survey was conducted to understand how drivers prefer to receive alerts during these incidents, which is essential for de-




(a) Eye gaze behavior during the NDRT phase




(b) Eye gaze behavior during the response task phase

**Fig. 4.** Main effect of NDRT engagement levels on eye gaze behavior. Each bar graph presents the mean and standard error for each factor level. Asterisks indicate significance in pairwise comparison. *$p < .05$. **$p < .01$. ***$p < .001$.

**Table 4**
The main effect of cyber-attack criticality levels and the interaction effects between NDRT engagement levels and cyber-attack criticality levels.

| | | Non-safety-related scenario | Safety-related scenario | Test Statistics | p-value |
|---|---|---|---|---|---|
| Perceived workload | Effort | 30.15 (23.06) | 33.04 (22.70) | $F(1, 132) = 4.10$ | $P = .044$ |

Note. Only statistically significant effects ($p < .05$) are presented.

signing effective cybersecurity interventions. Participants were asked to indicate their preferences for various IVIS-based alert types during cyber-attacks. The results showed that all participants, regardless of their NDRT engagement level, recognized the importance of multimodal alerts, including speech-based, visual, and combined alerts. Thirteen participants expressed a preference for speech alerts that not only indicate the presence of a cyber-attack but also provide an explanation of the situation and suggested actions. One participant noted, *"I often find myself occupied with other tasks, especially during longer driving routes, so having speech-based alerts would be beneficial. This way, I can divert my attention without needing to look at the visual display."*

Additionally, nine participants suggested that alert modality should be context-dependent, advocating speech alerts in safety-critical situations and visual alerts for non-critical scenarios to prevent cognitive overload. Thirteen participants preferred text-based visual alerts that incorporate bold fonts, pop-ups, and color coding to indicate the severity of an attack, ensuring that alerts are easily distinguishable from routine notifications. Furthermore, six participants advocated for heads-up display (HUD) alerts projected onto the windshield, as they believed this method would allow them to stay visually engaged with the road while processing cybersecurity warnings. As one participant explained, *"It would be beneficial if the alerts could be displayed on the windshield, allowing me to simultaneously view road information and the alert."* These findings align with the study's objective of investigating how task engagement and cyber-attack criticality influence driver response behavior. Since NDRT engagement can delay visual attention shifts and response times, alert systems must be designed to effectively redirect attention to cybersecurity threats without overwhelming the driver. The results suggest that multimodal alerts tailored to attack severity and driver task engagement levels may be the most effective strategy for ensuring timely and appropriate responses to cyber-attacks in automated vehicles.

## 5. Discussion

This study investigated the effects of Non-Driving Related Task (NDRT) engagement levels and cyber-attack criticality on drivers' situation awareness, eye-gaze behavior, response time, and perceived workload. The primary research objectives were:

1. To examine how varying levels of NDRT engagement influence driver responses to cyber-attacks, including attention allocation, situation awareness, response efficiency, and workload.
2. To assess how the criticality of cyber-attacks affects driver behavior and cognitive processing, particularly in safety-related vs. non-safety-related cyber-attacks.

The results revealed that Single NDRT and Dual NDRT engagement significantly reduced situation awareness, altered gaze pattern, increased response time, and elevated perceived workload, confirming H1a, H1b, H1c, and H1d. In contrast, cyber-attack criticality had a significant impact only on perceived workload (supporting H2d), while its effects on situation awareness, gaze behavior, and response time were not statistically significant.

### 5.1. Situation awareness (SA)

We hypothesized that Single NDRT and Dual NDRT engagement would result in lower situation awareness (H1a) due to limited cognitive resources available for monitoring the road environment (Wickens and Liu, 1988). The results support this assumption: participants in the Single and Dual NDRT groups exhibited significantly lower SA scores across all three SA levels compared to the No NDRT group (Table 3). These findings align with prior studies emphasizing the importance of SA in responding to automation failures. Clark et al. (2017) found that drivers with higher SA during automation failures exhibited quicker responses, supporting the idea that situation awareness is critical for timely interventions in AV cybersecurity incidents. Similarly, Merat et al. (2014) emphasized that drivers must maintain sufficient awareness of vehicle and roadway conditions to effectively manage automation failures. Consistent with this, our findings indicate that increased NDRT engagement significantly reduces SA, impairing a driver's ability to detect and respond to cyber-attacks in a timely manner. Interestingly, cyber-attack criticality did not significantly impact SA (H2a), suggesting that situation awareness was affected more by NDRT engagement than by the perceived severity of the attack. This finding implies that drivers may struggle to differentiate between safety-critical and non-safety-critical threats when re-engaging from NDRTs, underscoring the need for clearer cybersecurity alerts that help drivers assess the urgency of a given attack.

### 5.2. Eye gaze behavior

To examine attention allocation during cyber-attacks (H1b, H2b), we analyzed glance durations across key AOIs: the road, IVIS screen, dashboard, and NDRT interface.

#### 5.2.1. Road: compensation strategies

During the NDRT phase, both Single and Dual NDRT groups spent significantly more time fixating on the road compared to the No NDRT group. Interestingly, the Dual NDRT group exhibited significantly higher road fixation than the Single NDRT group, possibly reflecting an instinctive compensation strategy for divided attention. However, this increased road fixation did not translate into improved situation awareness, reinforcing previous findings on inattentional blindness (Simons and Chabris, 1999), where fixating on a visual stimulus does not necessarily equate to processing its meaning.

After the cyber-attack onset, Single NDRT group exhibited significantly greater road fixation than the No NDRT group, suggesting an attempt to reallocate attention back to the driving environment. However, the Dual NDRT group failed to significantly increase road fixation, indicating greater cognitive inertia and delayed attentional flexibility. These findings align with prior research on cognitive workload and attention shifts, where highly demanding tasks delay a driver's ability to redirect focus to driving-relevant stimuli after an event occurs (Wickens and Gutzwiller, 2017; Louw et al., 2019).

#### 5.2.2. IVIS screen: dependency on system-generated alerts

During the NDRT phase, Single NDRT group exhibited significantly greater IVIS screen fixation than the No NDRT group, while Dual NDRT group spent significantly less time looking at the IVIS than the Single NDRT group. This suggests that Single NDRT group were more engaged with the IVIS system, possibly due to anticipatory monitoring of cybersecurity alerts. In contrast, Dual NDRT group, who were juggling multiple secondary tasks, divided their attention further, leading to reduced IVIS engagement. Prior research has shown that drivers' engagement with in-vehicle screens is affected by task complexity, with higher cognitive loads limiting the ability to efficiently monitor secondary but important sources of information (Parker et al., 2022; Metz et al., 2011). Following a cyber-attack, Single NDRT group exhibited greater IVIS fix-

ation than the No NDRT group, suggesting anticipatory monitoring of cybersecurity alerts. However, Dual NDRT group spent significantly less time on the IVIS screen, indicating that increased cognitive load limited their ability to effectively monitor system-generated alerts.

Following a cyber-attack, IVIS fixation substantially increased in the Single NDRT group, suggesting greater reliance on visual cues for attack interpretation. However, Dual NDRT group failed to reallocate attention to the IVIS system, likely due to task disengagement delays. This finding is consistent with previous research showing that higher cognitive load can lead to attentional rigidity, limiting the ability to shift gaze between competing tasks (Harbluk et al., 2007; Louw et al., 2019).

#### 5.2.3. Dashboard: neglected vehicle status monitoring

For NDRT phase, both Single and Dual NDRT groups spent significantly less time looking at the dashboard compared to the No NDRT group. This suggests that once engaged in an NDRT, participants deprioritized monitoring essential vehicle status indicators, such as speed, and system warnings. Previous research suggests that drivers engaged in secondary tasks may neglect dashboard monitoring (Metz et al., 2011), reducing awareness of potential system failures.

#### 5.2.4. NDRT interface: prolonged engagement in NDRTs

As expected, Dual NDRT group spent significantly more time looking at the NDRT interface compared to both the Single NDRT and No NDRT groups, suggesting deep cognitive engagement in NDRTs. During the cyber-attack response phase, Dual NDRT group exhibited slower disengagement from the NDRT interface compared to Single NDRT group, indicating a lag in shifting visual attention back to critical information (Wickens and Gutzwiller, 2017). This suggests that multitasking in AVs may delay the ability to redirect attention, ultimately impairing cyber-attack detection and response.

These findings highlight the challenges of attentional flexibility in automated driving, particularly during cyber-attack incidents, emphasizing the need for adaptive in-vehicle interface designs that support rapid attention reallocation and intuitive cybersecurity alerting mechanisms.

### 5.3. Response time

Although Level 4 Avs are designed to function independently without human intervention in normal conditions, the ability of drivers to respond when unexpected cybersecurity threats arise remains a crucial concern. Cyber-attacks introduce unpredictable disruptions that may require immediate intervention, making response time a key indicator of driver readiness in highly automated environments. The results demonstrated that 100 % of participants pressed the brake pedal in response to cyber-attacks. The results showed that participants engaged in Single and Dual NDRTs exhibited significantly longer response times compared to those in the No NDRT group (Table 3). This finding indicates that NDRT engagement hinders the efficient reallocation of cognitive resources, delaying drivers' ability to detect and react to cyber-attacks. The prolonged response times observed in NDRT groups suggest that the process of disengaging from an NDRT, reorienting to the driving environment, and interpreting the cyber-attack alert imposes additional cognitive demands, contributing to slower reaction speeds. Notably, even participants in the No NDRT group exhibited an average response time of 3.79 s, which exceeds the 2.15 s average takeover time reported in a meta-analysis of Avs (Zhang B et al., 2019). This suggests that responding to cyber-attacks—even when not engaged in an NDRT—requires significant cognitive effort, likely due to the novel and ambiguous nature of cybersecurity threats. Unlike traditional takeover scenarios in Level 3 AVs, where system disengagement alerts are direct and clear, cyber-attacks may involve gradual or hidden manipulations, requiring drivers to recognize anomalies, assess risks, and determine an appropriate response—all of which take time.

### 5.4. Perceived workload

The findings indicate that NDRT engagement significantly impacted overall workload, mental demand, temporal demand, and perceived performance (Table 3). Participants in both Single and Dual NDRT groups reported higher overall workload during the response task to cyber-attacks compared to the No NDRT group, suggesting that engagement in NDRTs increased cognitive demands and effort when responding to cybersecurity threats.

However, mental demand was significantly higher only in the Single NDRT group compared to both the No NDRT and Dual NDRT groups. This distinction may be explained by differences in sensory modality utilization between tasks. According to Multiple Resource Theory (Wickens, 2002), tasks that rely on overlapping cognitive resources compete for attention, leading to increased cognitive strain. In contrast, the Dual NDRT group performed tasks that engaged separate sensory modalities (e.g., verbal vs. visual), potentially reducing interference and allowing for more efficient cognitive resource allocation. For example, a participant simultaneously engaging in conversation (auditory and verbal) while typing (visual and manual) may have experienced less interference than one performing only a visually demanding task.

In addition to NDRT engagement, cyber-attack criticality also significantly influenced workload (Table 4). Safety-related cyber-attacks resulted in significantly higher perceived effort compared to non-safety-related scenarios, likely due to the increased risk associated with system failures that could affect vehicle control and road safety. This highlights the critical importance of allocating sufficient cognitive resources to manage cybersecurity threats, particularly in high-risk situations that could lead to accidents or fatalities.

The interaction effect between NDRT engagement and cyber-attack criticality was also examined (Table 5). However, post-hoc pairwise comparisons did not reveal statistically significant differences, suggesting that while both factors contribute to workload, their combined effects may be more complex and require further investigation. In general, the heightened effort observed in safety-critical cyber-attacks may be attributed to the urgency and complexity of managing threats that directly impact vehicle dynamics, potential hazards, and risk mitigation strategies.

Table 6

Previous research suggests that workload transitions can lead to performance declines, particularly in high-risk contexts (Bowers et al., 2014; Cox-Fuenzalida, 2017; Cumming et al., 1973; Goldberg et al., 1980). In safety-related scenarios, shifting from a cognitively engaging NDRT to an urgent cybersecurity response task may exacerbate cognitive overload, delaying response execution and increasing error likelihood.

**Table 5**
Interaction effects between NDRT engagement levels and cyber-attack criticality levels.

| | | Non-safety-related scenario | Safety-related scenario | Test Statistics | p-value |
|---|---|---|---|---|---|
| Perceived workload Effort | **No NDRT Mean (SD)** | 30.77 (24.56) | 39.55 (26.08) | F(2, 132) = 3.62 | $P = .029$ |
| | **Single NDRT Mean (SD** | 29.55 (20.36) | 42.88 (24.85) | | |
| | **Dual NDRT Mean (SD)** | 30.11 (24.50) | 45.11 (30.72) | | |

Note. Only statistically significant effects ($p < .05$) are presented.

**Table 6**
A post-experiment survey questions.

| Short survey on designing a display interface |
|---|
| 1 How would you redesign a display interface, be it a voice-agent or a visual display, to enhance your preparedness for unforeseen cyber-security threats? *For example, this could involve incorporating voice agents that provide explanations of the situation and guidance on resolving it.* If so, what kinds of alerts would you like to have? |
| 2 How would you redesign a display interface according to different types of cyber security scenarios?*For example, you may perceive some sort of attack scenarios are more critical than the others in terms of safety/threat.* • Owner name of phone has changed • Final destination has changed • Speed has changed • Old data is repeated • ETA has changed • Out of service If so, what kinds of alerts would you like to have for which attack scenarios? |

### 5.5. Design implications

Based on the results of the current study, including survey question results, we propose several interface design implications to enhance driver interaction with AV cybersecurity systems during cyber-attacks.

(1) Minimize Visual Attention Disruptions to Enhance Situation awareness.The results indicate that drivers engaged in Non-Driving Related Tasks (NDRTs) demonstrated delayed reallocation of visual attention to critical driving information during cyber-attacks. Prolonged IVIS fixation among Single NDRT group and excessive NDRT interface fixation among Dual NDRT group suggest that in-vehicle interface designs must minimize excessive glance durations away from the road. For example, place critical cyber-attack alerts within the driver's primary field of view (e.g., using Head-Up Displays (HUDs) or Augmented Reality overlays) to minimize off-road glances.

(2) Clearly Differentiate Safety-Critical vs. Non-Safety-Critical Cyber-Attack Alerts.The study found that safety-related cyber-attacks imposed significantly higher cognitive demands compared to non-safety-related attacks. However, delayed driver responses suggest that distinguishing between different types of cybersecurity threats was not always intuitive. For example, use distinct color coding and iconography to categorize safety-related vs. non-safety-related cyber-attacks (e.g., red for critical alerts, yellow for moderate threats, and blue for informational messages), and also integrate multimodal alerting mechanisms, such as haptic seat vibrations for critical threats and verbal cues for lower-priority notifications, ensuring that drivers can quickly interpret the urgency of a cyber-attack without excessive cognitive processing.

(3) Provide Contextual Information to Improve Cyber-Attack Comprehension.The results revealed that longer response times were observed even in drivers not engaged in NDRTs, suggesting that cyber-attacks may not always be immediately understood. Unfamiliarity with cybersecurity threats may require additional cognitive processing time, delaying effective responses. To mitigate these delays, AV cyber-attack alerts should provide contextual information that enhances drivers' understanding of the attack and its potential consequences. For example, use in-vehicle conversational agents or verbal prompts to summarize attack type, affected vehicle systems, and recommended actions (e.g., "Unauthorized GPS spoofing detected. Your vehicle location may be inaccurate. Reconfirm your route.").

(4) Personalize Cybersecurity Alerts to Align with Individual Driving and Risk Preferences.The results revealed that drivers themselves expressed a desire for customizable alert settings. For instance, three participants suggested a "High alert mode" delivering frequent updates, whereas another preferred a "Minimal alert mode" that only presents critical notifications. These comments align with the observed variability in gaze behavior and response strategies across NDRT engagement levels, indicating that a one-size-fits-all alerting system may not be optimal. Accordingly, AV interfaces should allow drivers to tailor both alert frequency and intensity based on their personal preferences and experience. Future work will quantify how different customization options affect performance and satisfaction.

## 6. Limitations and future work

Limitations of the current study are described here along with possible future research directions: first, our experiment limited non-driving-related task (NDRT) engagement to approximately 2 min. In naturalistic Level 4 AV use, drivers may remain disengaged for much longer, potentially leading to increased cognitive disengagement or drowsiness (Gonçalves et al., 2015). Future research should examine how prolonged NDRT engagement (e.g., 15–30 min or more) affects cognitive workload, situation awareness, and response behavior during cybersecurity incidents, and how fatigue interacts with attack-response performance. Second, we modeled cyber-attacks exclusively through perceptual IVIS alert manipulations for experimental control, and confined participant responses to a single brake-pedal action. While this approach allowed us to isolate alert detection and reaction time, it does neither replicate the urgency of attacks that directly alter vehicle behavior (e.g., unintended acceleration, braking failures) nor capture the range of real-world takeover strategies. Future studies should therefore (a) integrate cyber-attacks that produce tangible vehicle responses—such as forced deceleration or lane deviations—and (b) offer multiple, open-ended response measures (e.g., time to first fixation, steering override, manual throttle control, IVIS confirmation taps or "What would you do first?" prompts) within the simulator. Third, our study employed only six discrete attack messages. The AV threat landscape spans diverse vectors and severities. Future work should broaden scenario variety—varying attack type (sensor spoofing vs. control-system takeover), severity, and timing—to map how different profiles influence detection, comprehension, and action. Fourth, we recruited only young drivers. Given age-related changes in attention and reaction time (Bhise, 2011) and the role of driving experience in NDRT engagement and response strategies (Mourant and Rockwell, 1972), future research should include older adults and treat driver experience as a moderating factor. Fifth, we did not measure participants' trust in the AV or IVIS alerts following cyber-attack messages. Since trust in automation can critically mediate takeover behavior, future studies should incorporate validated trust scales via post-drive surveys or separate freeze-period probes, and analyze how trust dynamics interact with attention, workload, and response efficacy under attack. Sixth, all participants experienced the same order of six attack onsets. While random assignment to NDRT groups balanced practice and fatigue effects, fixed timing may still permit temporal anticipation. Subsequent experiments should introduce variability or counterbalance attack timings (e.g., Latin-square or randomized schedules) to assess how unpredictability influences detection and response. Finally, participants were not asked whether they could discern "safety-related" versus "non-safety-related" messages or identify which of the six scenarios they experienced. As a result, we cannot determine how explicit awareness affects detection speed or accuracy. Future work should include recognition checks or confidence ratings to examine the interplay between conscious scenario knowledge, NDRT engagement, and performance.

## 7. Conclusion

This study revealed that varying levels of NDRT engagement significantly impact drivers' situation awareness regarding road information, eye-gaze behavior, response time, and perceived workload. Participants engaged in NDRTs exhibited reduced road situation awareness, indicating that their ability to perceive and interpret critical roadway information was diminished during cyber-attacks. This suggests that prolonged engagement in NDRTs may impair a driver's readiness to detect and respond to cybersecurity threats, as they are less attuned to vehicle behavior and environmental cues when an attack occurs. Additionally, cyber-attack criticality levels influenced drivers' perceived workload, with safety-related attacks eliciting significantly higher cognitive effort compared to non-safety-related threats. This suggests that drivers allocate greater mental resources to processing cybersecurity incidents that pose immediate risks to vehicle control and road safety. The increased workload associated with these events reinforces the need for intuitive and adaptive AV interfaces that support efficient decision-making and threat mitigation, particularly in high-risk scenarios.

From a human-machine interface (HMI) and cybersecurity design perspective, this study offers valuable insights into optimizing in-vehicle interfaces to enhance driver response during cyber-attacks. Minimizing unnecessary visual engagement, prioritizing critical safety information, implementing adaptive alert mechanisms, and providing clear contextual information can help drivers effectively interpret and respond to cybersecurity threats while maintaining situation awareness. Designing multi-modal alerting systems that integrate auditory, haptic, and visual cues can further support efficient workload management and reduce cognitive overload in cybersecurity-critical situations.

While prior research has explored cybersecurity risks in automated vehicles, this study represents a unique contribution by investigating how NDRT engagement and cyber-attack criticality influence drivers' situation awareness of road information, visual attention, response time, and workload in Level 4 AVs. By examining how drivers allocate attention and execute responses under cyber-attack conditions, this study provides a basis for future work to strengthen cybersecurity resilience, improve driver safety, and refine AV interface design.

## Uncited references

Lansdown, 2000, Ahmad et al., 2020, Blindness, 2010, Cox-Fuenzalida, 2007.

## CRediT authorship contribution statement

**Gayoung Ban:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Myounghoon Jeon:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijhcs.2025.103554.

## References


Aliebrahimi, S., Miller, E.E., 2023. Effects of cybersecurity knowledge and situation awareness during cyberattacks on autonomous vehicles. Transp. Res. Part f: Traffic Psychol. Behav. 96, 82–91.

Andrade, R.O., Yoo, S.G., 2019. Cognitive security: a comprehensive study of cognitive science in cybersecurity. J. Inf. Secur. Appl. 48, 102352. https://doi.org/10.1016/j.jisa.2019.06.008.

Ahmad, U., Song, H., Bilal, A., Alazab, M., Jolfaei, A., 2020. Securing smart vehicles from relay attacks using machine learning. J. Supercomput. 76, 2665–2682.

Bansal, P., Kockelman, K.M., 2017. Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies. Transp. Res. A: Policy Pract. 95, 49–63.

Bowers, M.A., Christensen, J.C., Eggemeier, F.T., 2014. The effects of workload transitions in a multitasking environment. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58. SAGE Publications, Sage CA: Los Angeles, CA, pp. 220–224.

Blindness, S.I, 2010. Gorillas in our midst: sustained inattentional blindness for dynamic events. In: Simons, Daniel J., Chabris, Christopher F. (Eds.), Perception: Visual perception. Perception: Visual perception, 2, p. 277.

Bhise, V.D., 2011. Ergonom Icsin the Autom Otive Design Process. CRC Press, Boca Raton.

Cox-Fuenzalida, L.E., 2007. Effect of workload history on task performance. Hum. Factors 49 (2), 277–291.

Cumming, R.W., Croft, P.G., 1973. Human information processing under varying task demand. Ergonomics 16 (5), 581–586.

Clark, H., McLaughlin, A.C., Feng, J., 2017. Situation awareness and time to takeover: exploring an alternative method to measure engagement with high-level automation. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61. SAGE Publications, Sage CA: Los Angeles, CA, pp. 1452–1456.

... &Dong, C., Chen, Y., Wang, H., Wang, L., Li, Y., Ni, D., Hua, X., 2024. Evaluating impact of remote-access cyber-attack on lane changes for connected automated vehicles. Digit. Commun. Netw. 10 (5), 1480–1492.

Eiza, M.H., Ni, Q., 2017. Driving with sharks: rethinking connected vehicles with vehicle cybersecurity. IEEE Veh. Technol. Mag. 12 (2), 45–51.

Endsley, M.R., 2000. Situation Awareness Analysis and Measurement. Lawrence Erlbaum Associates.

Goldberg, R.A., Stewart, M.R., 1980. Memory overload or expectancy effect?'Hysteresis' revisited. Ergonomics 23 (12), 1173–1178.

Golbabaei, F., Yigitcanlar, T., Bunker, J., 2021. The role of shared autonomous vehicle systems in delivering smart urban mobility: a systematic review of the literature. Int. J. Sustain. Transp. 15 (10), 731–748.

... &Gonçalves, M., Amici, R., Lucas, R., Åkerstedt, T., Cirignotta, F., Horne, J., Aksu, M., 2015. Sleepiness at the wheel across Europe: a survey of 19 countries. J. Sleep. Res. 24 (3), 242–253.

Greenblatt, J.B., Shaheen, S., 2015. Automated vehicles, on-demand mobility, and environmental impacts. Curr. sustain./renew. energy rep. 2, 74–81.

Harbluk, J.L., Noy, Y.I., Trbovich, P.L., Eizenman, M., 2007. An on-road assessment of cognitive distraction: impacts on drivers' visual behavior and braking performance. Accid. Anal. Prev. 39 (2), 372–379.

He, F., Elhammady, K., Fischmeister, S., Burns, C.M., 2024. Preliminary cognitive modeling: comparing distraction-based cyber-attacks and alcohol-related impairments on Human drivers. In: 2024 IEEE 4th International Conference on Human-Machine Systems (ICHMS). IEEE, pp. 1–6.

Hart, S.G., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Hum. Ment. Workload/Elsevier.

Ko, S.M., Ji, Y.G., 2018. How we can measure the non-driving-task engagement in automated driving: comparing flow experience and workload. Appl. Erg. 67, 237–245.

Lansdown, T.C., 2000. Driver visual allocation and the introduction of intelligent transport systems. Proc. Inst. Mech. Eng. D: J. Automob. Eng. 214 (6), 645–652.

Louw, T., Kuo, J., Romano, R., Radhakrishnan, V., Lenné, M.G., Merat, N., 2019. Engaging in NDRTs affects drivers' responses and glance patterns after silent automation failures. Transp. Res. Part f: Traffic Psychol. Behav. 62, 870–882.

Lim, C., Prendez, D., Boyle, L.N., Rajivan, P., 2024. The impact of cybersecurity attacks on Human trust in autonomous vehicle operations. Hum. Factors 00187208241283321.

Marcinkiewicz, V., Zhang, Q., & Morgan, P. (2023). The effects of cyber readiness and response on human trust in self driving cars.

Metz, B., Schömig, N., Krüger, H.P., 2011. Attention during visual secondary tasks in driving: adaptation to the demands of the driving task. Transp. Res. Part f: Traffic Psychol. Behav. 14 (5), 369–380.

Meryem, S., Mazri, T., 2019. Security study and challenges of connected autonomous vehicles. In: Proceedings of the 4th International Conference on Smart City Applications. pp. 1–4.

Merat, N., Jamson, A.H., Lai, F.C., Daly, M., Carsten, O.M., 2014. Transition to manual: driver behaviour when resuming control from a highly automated vehicle. Transp. Res. Part F: Traffic Psychol. Behav. 27, 274–282.

Mourant, R.R., Rockwell, T.H., 1972. Strategies of visual search by novice and experienced drivers. Hum. Factors 14 (4), 325–335.

Noy, I.Y., Shinar, D., Horrey, W.J., 2018. Automated driving: safety blind spots. Saf. Sci. 102, 68–78.

Nieuwenhuijsen, J., de Almeida Correia, G.H., Milakis, D., Van Arem, B., van Daalen, E., 2018. Towards a quantitative method to analyze the long-term innovation diffusion of automated vehicles technology using system dynamics. Transp. Res. C: Emerg. Technol. 86, 300–327.

Payre, W., Perelló-March, J., Sriranga, A.K., Birrell, S., 2023. The notorious BIT: the

effects of a ransomware and a screen failure on distraction in automated driving. Transp. Res. Part f: traffic Psychol. Behav. 94, 42–52.
Parker, J.I., Zhang, F., Wang, M., Roberts, S.C., 2022. How do drivers respond to vehicle cyberattacks? A driving simulator study. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 66. SAGE Publications, Sage CA: Los Angeles, CA, pp. 737–741.
Parkinson, S., Ward, P., Wilson, K., Miller, J., 2017. Cyber threats facing autonomous and connected vehicles: future challenges. IEEe trans. Intell. Transp. Syst. 18 (11), 2898–2915.
... &Pendleton, S.D., Andersen, H., Du, X., Shen, X., Meghjani, M., Eng, Y.H., Ang, M.H., 2017. Perception, planning, control, and coordination for autonomous vehicles. Machines 5 (1), 6.
Petit, B.J., Stottelaar, B., Feiri, M., & Kargl, F. (2015). Remote attacks on automated vehicles sensors: experiments on camera and LiDAR black Hat Europe. Amsterdam, Netherlands.
Pyke, A., Bouchelle, R., Uzhca, D., 2023. Out of sight but still In mind: making 'invisible' Cyber threats more salient via concrete analogies. Hum. Factors Cybersecur. (91), 91.
Rofail, M., Alsafty, A., Matousek, M., Kargl, F., 2019. Multi-modal deep learning for vehicle sensor data abstraction and attack detection. In: 2019 IEEE International Conference on Vehicular Electronics and Safety (ICVES). IEEE, pp. 1–7.
SAE (2016) Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles (Surface Vehicle Recommended Practice: superseding J3016 Jan 2014), SAE International, September 2016
Siddiqui, F., Khan, R., Tasdemir, S.Y., Hui, H., Sonigara, B., Sezer, S., McLaughlin, K., 2023. Cybersecurity engineering: bridging the security gaps in advanced automotive systems and ISO/SAE 21434. In: 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring). IEEE, pp. 1–6.
Simons, D.J., Chabris, C.F., 1999. Gorillas in our midst: sustained inattentional blindness for dynamic events. Perception. 28 (9), 1059–1074.
Sheehan, B., Murphy, F., Ryan, C., Mullins, M., Liu, H.Y., 2017. Semi-autonomous vehicle motor insurance: a Bayesian Network risk transfer approach. Transp. Res. C: Emerg. Technol. 82, 124–137.
Thing, V.L., Wu, J., 2016. Autonomous vehicle security: a taxonomy of attacks and defences. In: 2016 ieee international conference on internet of things (ithings) and ieee green computing and communications (greencom) and ieee cyber, physical and social computing (cpscom) and ieee smart data (smartdata). IEEE, pp. 164–170.
Wang, N.W., Huang, Y.M., Chen, W.M., 2008. A novel secure communication scheme in vehicular ad hoc networks. Comput. Commun. 31 (12), 2827–2837.
Wang, M., Zhang, F., Roberts, S.C., 2024. A simulator study assessing the effectiveness of training and warning systems on drivers' response performance to vehicle cyberattacks. Accid. Anal. Prev. 203, 107644.
Wickens, C.D., Liu, Y., 1988. Codes and modalities in multiple resources: a success and a qualification. Hum. Factors 30 (5), 599–616.
Wickens, C.D., 2002. Multiple resources and performance prediction. Theor. Issues. Ergon. Sci. 3 (2), 159–177.
Wickens, C.D., Gutzwiller, R.S., 2017. The status of the strategic task overload model (STOM) for predicting multi-task management. In: Proceedings of the human factors and ergonomics society annual meeting, 61. SAGE Publications, Sage CA: Los Angeles, CA, pp. 757–761.
Yao, F., Zhu, J., Yu, J., Chen, C., Chen, X.M., 2020. Hybrid operations of human driving vehicles and automated vehicles with data-driven agent-based simulation. Transp. Res. D: Transp. Environ. 86, 102469.
Zhou, Y., Xu, M., 2023. Robotaxi service: the transition and governance investigation in China. Res. Transp. Econ. 100, 101326.
Zhang, B., De Winter, J., Varotto, S., Happee, R., Martens, M., 2019a. Determinants of take-over time from automated driving: a meta-analysis of 129 studies. Transp. Res. Part F: Traffic Psychol. Behav. 64, 285–307.
Zhang, F., Petit, J., Roberts, S.C., 2019b. A simulator study on drivers' response and perception towards vehicle cyberattacks. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 63. SAGE Publications, Sage CA: Los Angeles, CA, pp. 1498–1502.
Zhang, F., Wang, M., Parker, J.I., Roberts, S.C., 2023. The effect of driving style on responses to unexpected vehicle cyberattacks. Safety 9 (1), 5.