



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/245436/>

Version: Accepted Version

Article:

Zhang, W., Carsten, O., Lee, Y.T. et al. (2026) Investigation of driving safety in partial driving automation considering driving modes and drivers' pre-drive states. Accident Analysis & Prevention, 236. 108664. 108664-. ISSN: 0001-4575 (In Press)

<https://doi.org/10.1016/j.aap.2026.108664>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Investigation of Driving Safety in Partial Driving Automation Considering Driving Modes and Drivers' Pre-drive States

Wei Zhang ¹, Oliver Carsten ², Yee Thung Lee ², Jinlei Shi ³, Wei Zhang ^{4*}

1 Department of Industrial Engineering, Tsinghua University, China

2 Institute for Transport Studies, University of Leeds, United Kingdom

3 School of Design, Southwest Jiaotong University, Chengdu, China

4 Professor, Department of Industrial Engineering, Tsinghua University, China

* Corresponding author.

Email: zhangwei@mail.tsinghua.edu.cn

Abstract

With the release of partial driving automation systems that allow drivers to take their hands off the steering wheel (L2 hands-off), it has become increasingly important to evaluate their implications for driving safety. This study investigated how different driving modes influence driving safety considering drivers' pre-drive states. Using a static driving simulator, two groups of UK drivers participated in the experiment: 19 non-nightshift workers and 19 nightshift workers. Participants in each group completed drives under three driving modes: manual driving, L2 hands-on, and L2 hands-off. The drowsiness, attention behavior, hazard perception performance, and driving performance of the participants were measured. The results suggested that the L2 hands-off mode had adverse effects on drowsiness and attention behaviors in both groups, with detrimental trends also observed for hazard perception performance and driving performance. Moreover, given the adverse effects of the hands-off mode, drivers with different pre-drive states adopted different response strategies. Non-nightshift workers performed more frequent takeovers in the hands-off mode to avoid collisions, whereas nightshift workers responded similarly across all modes but exhibited the highest collision probability in the hands-off mode. In summary, although L2 driving may reduce manual control demands, taking hands off steering wheel increases safety risks, particularly for nightshift workers. Future efforts should adopt strategies to mitigate the safety risks associated with L2 hands-off systems and assess drivers' pre-drive states before permitting the use of such systems.

Keywords: L2 driving system; Drowsiness; Hazard perception performance; Hands-on mode; Hands-off mode

1 Introduction

1.1 Background

Partial driving automation (i.e., Level 2 [L2] driving), which can perform both steering and longitudinal control (SAE International, 2021), has been integrated into transportation systems and is now widely adopted. Unlike manual driving, which keeps drivers “in the loop” through continuous physical control, L2 systems do not require such ongoing physical control but instead require drivers to remain “on the loop” by continuously monitoring the driving environment (Merat et al., 2019). In early stages, L2 systems that required drivers to keep their hands on steering wheel were predominant on the market. More recently, systems such as Ford’s BlueCruise have been introduced in the UK, allowing drivers to move their hands away from the steering wheel under specific conditions (European Transport Safety Council, 2023).

Although L2 hands-off systems may be appealing to drivers who prefer reduced hand engagement (Gershon et al., 2023), the shift from “hands-on” to “hands-off” may alter drivers’ states, attentional behaviors, and driving responses (Lee et al., 2024; Pan, Logan, et al., 2025). As a result, L2 hands-off systems may further increase the likelihood of drivers entering an “out-of-the-loop” state, in which they become disengaged from both vehicle control and environmental monitoring (Merat et al., 2019). Therefore, understanding how hand engagement influences driving safety is critical for the safe deployment of L2 hands-off systems. This study examines the potential effects of L2 hands-off systems on driving safety from four perspectives: driver drowsiness, attention behaviors, hazard perception performance, and driving performance.

1.2 Influence of hand engagement

1.2.1 Driver drowsiness

Changing hand engagement from “hands-on” to “hands-off” in L2 driving might exacerbate driver drowsiness, which is a state of reduced alertness that contributes to approximately 20% of motor vehicle crashes (National Sleep Foundation, 2023).

Drowsiness can impair drivers' performance by reducing executive functioning, decreasing attention to the road environment, slowing responses to hazards, and consequently increasing crash risk (American Psychological Association, 2018; Kaplan & Prato, 2012; Smith et al., 2009). In autonomous driving, driver drowsiness mainly arises from task underload (May & Baldwin, 2009) and tends to increase with both the duration of automated driving and the level of automation, especially when no non-driving-related task (NDRT) is present (Boguslavsky et al., 2025; Pan, Logan, et al., 2025; Vogelpohl et al., 2019). Compared with hands-on mode, taking hands off the steering wheel removes drivers' primary physical contact with the driving task. Since physical effort has been shown to increase vigilance (Smit et al., 2005), this reduction in hand engagement may further increase task underload and thereby promote drowsiness (May & Baldwin, 2009). In contrast, keeping hands on the steering wheel may help drivers remain physically involved in the driving task. This may potentially mitigate underload-related drowsiness.

How hand engagement influences driver drowsiness remains unclear. Although previous studies have examined the effects of automation on driver drowsiness (e.g., Boguslavsky et al., 2025; Greenlee et al., 2018; Körber et al., 2015; Pan, Logan, et al., 2025; Vogelpohl et al., 2019), they did not focus on hand engagement. For example, in a L2 driving experiment, Körber et al. (2015) found that driver drowsiness increased significantly over time. Vogelpohl et al. (2019) reported that drivers became drowsy earlier and more severely in autonomous driving than in manual driving. Although Soulmana et al. (2021) considered both hand positions and drowsiness, their study focused on how drivers' hand positions and drowsiness changed during driving, rather than evaluating the effect of hand engagement on driver drowsiness in L2 driving. Therefore, the influence of hand engagement in driver drowsiness during L2 driving requires further investigation.

1.2.2 Attention behaviors and hazard perception performance

Changing hand engagement from “hands-on” to “hands-off” in L2 driving may also alter drivers’ attention behaviors. When drivers’ hands are free from the driving task in L2 driving, they may be more likely to divert their attention away from the driving environment to NDRTs (Carsten et al., 2012; Gershon et al., 2023; Hussain et al., 2025; Noble et al., 2021). Although attention shifts can also occur in the hands-on mode, maintaining physical contact with the steering wheel helps drivers remain more engaged in the driving task and thereby allocate more attention to road conditions. The influence of hand engagement on attention behavior has been partly supported by Lee et al. (2024), who found that longer hands-off durations were associated with shorter on-road glance times.

Moreover, the greater diversion of attention during L2 hands-off driving may further impair drivers’ hazard perception, which refers to their ability to detect dangerous traffic situations (Horswill & McKenna, 2004). Hazard perception involves scanning the traffic environment, allocating attention to critical stimuli, and forming a holistic interpretation of hazard salience (Milech et al., 1989; Smith et al., 2009). The hazard perception performance is often measured using drivers’ response behaviors (e.g., response reaction time and correct detections) (Greenlee et al., 2018; Huestegge et al., 2010) and eye-tracking measures (e.g., hazard identification, time to first fixation, etc.) (Borowsky et al., 2012; Cao et al., 2022; Crundall et al., 2012; Yang et al., 2021; Zangi et al., 2022). When visual attention is diverted from the roadway to NDRTs, drivers are less likely to detect potential hazards (Zangi et al., 2022). Therefore, compared with L2 hands-on mode, the greater shift of attention away from the driving environment in the hands-off mode may further impair hazard perception performance.

Drivers must detect and respond to events that exceed system capabilities (Cao et al., 2022). Understanding drivers’ attention behaviors and hazard perception performance is critical for L2 driving safety, especially under hands-off conditions. However, how hand engagement in L2 driving influences drivers’ attention behaviors

and hazard perception performance remains unclear. Although some studies have examined drivers' attention behaviors related to hand engagement in L2 driving (Gershon et al., 2023; Lee et al., 2024), they have mainly focused on associations between drivers' hand positions and attention in real-world L2 driving, rather than directly testing the effect of hand engagement on attention behaviors. On the other hand, although hazard perception performance in L2 driving has been investigated (Boguslavsky et al., 2025; Greenlee et al., 2018, 2024; Samuel et al., 2020; Zangi et al., 2022), most of previous studies were focused on L2 hands-on systems. For instance, the literature has found that drivers' hazard perception performance using hands-on mode was inferior to that using manual driving (Samuel et al., 2020). This performance also decreased over time in L2 driving (Greenlee et al., 2024). While Lee et al. (2025) considered the effects of L2 hands-off mode on hazard perception performance and compared it with manual driving, they did not compare hazard perception performance across different hand engagement modes. Therefore, it is necessary to further examine how hand engagement influences drivers' attention behaviors and hazard perception performance.

1.2.3 Driving performance

The state changes and attentional shifts described above may further lead to deficits in driving performance in critical situations that require drivers to switch back to manual control. Because drowsiness impairs drivers' executive functioning and reaction time (Foong et al., 2015), the greater tendency toward drowsiness in hands-off driving may further slow drivers' responses to hazards. On the other hand, as attentional shifts associated with NDRTs can adversely affect driving performance by reducing takeover frequency in L2 driving (Zangi et al., 2022), drivers in the hands-off mode may show poorer responses to hazards because they are more likely to divert their attention away from the road (Lee et al., 2024).

Compared with the other safety-related aspects, relatively more studies have examined the influence of hand engagement on driving performance (Cahour et al.,

2021; Damböck et al., 2013; Gershon et al., 2023; Hussain et al., 2025; Lee et al., 2025; Naujoks et al., 2015). Most of these studies have reported adverse effects of hands-off mode on driving performance, including slower steering responses (Hussain et al., 2025), longer takeover times (Gershon et al., 2023), and greater vehicle trajectory deviations (Cahour et al., 2021), compared with hands-on mode. However, some studies did not find significant differences between hands-on and hands-off modes in response behaviors (Damböck et al., 2013; Naujoks et al., 2015). Damböck et al. (2013), for example, did not find significant differences in driving behaviors between L2 hands-on and -off modes even though they did find late and longer takeover response of these modes than the manual driving. While Naujoks et al. (2015) also found no significant differences between the hands-on and hands-off modes, they argued that this was because many drivers maintained hand contact with the steering wheel even in the hands-off mode. This underscores the potential importance of maintaining hand contact with the steering wheel for driving performance. In summary, although the effects of hand engagement in L2 driving on driving performance have been partly explored, how hand engagement influences driving performance needs further investigation.

1.3 Drivers' pre-drive state

In addition, drivers' pre-drive states (e.g., alert or drowsy) play a critical role in road safety and should be emphasized in driving research (Radlmayr et al., 2019). Many studies have examined the influence of different driving modes under the assumption that drivers are alert (e.g., Körber et al., 2015; Lee et al., 2025; Zangi et al., 2022). Studies incorporating drivers who begin driving in a drowsy state are less frequent. Pre-drive drowsiness is common among groups such as nightshift workers and long-haul drivers. It may further aggravate driver drowsiness during autonomous driving (Ahlström et al., 2021) and impair hazard perception performance (Smith et al., 2009; Watling & Home, 2022). Considering drivers' pre-drive states when evaluating the effects of hand engagement in L2 driving is, therefore, an important road safety issue.

The influence of hand engagement on drowsiness, attention behaviors, hazard perception performance, and driving performance may differ between drowsy and alert drivers. The potential adverse effects of the hands-off mode may be more pronounced among drivers who are already drowsy. This is because the detrimental effect of hands-off driving may be exacerbated when combined with sleep restriction, leading drowsy drivers to experience even greater sleepiness. This heightened drowsiness during hands-off driving may, in turn, further impair hazard perception performance (Cao et al., 2022) and driving performance (Vogelpohl et al., 2019).

It remains unclear how different hand engagement modes in L2 driving influence drowsiness, attention behaviors, hazard perception performance, and driving performance across driver states. Some studies have shown that drivers in alert and drowsy states responded differently to changes in driving modes (Ahlström et al., 2021; Vogelpohl et al., 2019; Watling & Home, 2022). For example, Ahlström et al. (2021) found that under sleep-deprived conditions, L2 driving increased drowsiness relative to manual driving. Such a decrease was not found when the participants were in alert state. Similar findings have been reported by Vogelpohl et al. (2019). Moreover, Watling and Home (2022) found that drivers' hazard perception performance declined with prolonged driving duration, and this effect was more pronounced under sleep-deprived conditions. Although these studies considered different driver states, they did not examine the role of hand engagement or provide a comprehensive assessment across multiple safety-related outcomes. Boguslavsky et al. (2025), on the other hand, have studied the effects of L2 driving on underload, drowsiness, attention behaviors, and hazard perception performance, they did not consider either hand engagement or pre-drive states. Therefore, the influence of hand engagement on driver drowsiness, attention behaviors, hazard perception performance, and driving safety under different pre-drive states requires further investigation.

1.4 Aim of this study

There is still limited research on how different driving modes influence drivers' drowsiness, attention behaviors, hazard perception performance, and driving performance even with the deployment of the L2 hands-off systems (European Transport Safety Council, 2023). Moreover, because drivers may differ in their pre-drive states, few studies have examined whether the effects of driving modes on driving safety vary across different pre-drive states. Therefore, this study had two aims: (1) to investigate how different driving modes (i.e., L2 hands-on, L2 hands-off, and manual driving) affect driving safety; and (2) to explore whether these effects vary depending on drivers' pre-drive states (i.e., alert or drowsy). Based on these aims, the following research questions (RQs) and hypotheses (Hs) were proposed:

RQ1: How do different driving modes influence driving safety?

H1: Relative to manual driving, L2 hands-off driving would have more detrimental effects on driving safety. Specifically:

H1.1 L2 hands-off driving leads to greater post-drive drowsiness.

H1.2 L2 hands-off driving leads to greater diversion of attention away from the road.

H1.3 L2 hands-off driving leads to poorer hazard perception performance.

H1.4 L2 hands-off driving results in poorer driving performance.

RQ2: How do the effects of different driving modes on driving safety vary across drivers' pre-drive states?

H2: Compared with alert drivers, the adverse effects of L2 hands-off driving on driving safety are more pronounced in drowsy drivers.

H2.1 The effect of L2 hands-off driving on increasing drowsiness are more pronounced in drowsy drivers than in alert drivers.

H2.2 The effect of L2 hands-off driving on diverting attention away from the road are more pronounced in drowsy drivers than in alert drivers.

H2.3 The effect of L2 hands-off driving on impairing hazard perception performance are more pronounced in drowsy drivers than in alert drivers.

H2.4 The effect of L2 hands-off driving on worsening driving performance are more pronounced in drowsy drivers than in alert drivers.

2 Method

2.1 Participants

This study was approved by the University of Leeds Research Ethics Committee (Reference code: 2024 2253-2757). This study recruited two groups of drivers based on their pre-drive condition: 1) the “alert” group, consisting of non-nightshift workers who participated during daytime hours; 2) the “drowsy” group, consisting of nightshift workers who participated immediately after completing a night shift.

For inclusion, participants were required to meet the following criteria: 1) have held a valid UK drivers’ license for at least two years; 2) be regular drivers, driving at least once per week; 3) have an annual mileage of at least 500 miles; 4) be aged between 22-65 years; 5) have normal or corrected to normal vision; 6) not be pregnant; 7) not take cardioactive medications, such as antidepressants, antipsychotics, or antihypertensives; 8) have no severe motion sickness, disorientation tendencies, or serious mobility issues in back, knees, or hips; 9) have no sleep disorders. Moreover, all night-shift workers needed to meet two additional requirements: 10) work at least three hours between 11:00 PM and 6:00 AM; and 11) refrain from caffeine, mints, teas, alcohol, or related beverages, smoking, or having a nap for at least 12 hours prior to the experiment. Participants who failed to meet any of the above criteria were excluded.

A priori power analysis indicated that a minimum of 36 participants was required. This estimation was calculated using the G*Power software (3.1.9.6) with an effect size of 0.25, an α of 0.05, a power of 0.90, the number of groups of 2, the number of measurements of 3, a correlation of 0.5 among repeated measures, and a non-sphericity correction of 1.

A total of 38 participants, including 19 night-shift workers (11 male, 8 female) and 19 non-nightshift workers (10 male, 9 female), joined. The Epworth Sleepiness Scale

(ESS) was adopted to measure the daytime sleepiness tendency for each participant before the trial (Johns, 1991). The demographic characteristics of the two groups are summarized in Table 1. No significant differences were found between the two groups in terms of age, driving years, annual mileage, and daytime sleepiness tendency.

Table 1. Demographics of two driver groups

	Non-nightshift workers			Night-shift workers		
	Range	Mean	SD	Range	Mean	SD
Age	27-61	42.158	9.287	22-64	34.947	13.468
Driving years	3-40	19.474	10.019	2-47	14.039	13.898
Annual mileage	4000-18000	7421.053	3641.059	700-30000	11194.737	7921.032
Daytime sleepiness tendency	2-11	6.316	2.668	1-15	6.579	3.687

Note. All four comparisons between the two groups were non-significant (all $p > 0.05$).

2.2 Apparatus and materials

2.2.1 Driving simulator

The experiment was conducted using a static driving simulator at the University of Leeds. The simulator consisted of a 49-inch curved LED monitor (5120×1440 pixels), a 7-inch LED monitor (1920×1080 pixels), a driving system (Fanatec, Landshut, Germany), an adjustable seat (Playseat® B.V., Doetinchem, the Netherlands), and an audio system (see Figure 1). The 49- and 7-inch monitor were used to present the driving scenarios and the instrument panel, respectively. The driving system included a force-feedback steering wheel and a pedal set. The audio system was used to display the driving sound. An iPad (9th generation; Apple Inc., Cupertino, CA, USA) was positioned to the left of the steering wheel and served as the infotainment display for Apple Music and Google Maps. Participants' eye movements were recorded using Tobii Glasses 2 (Tobii AB (publ), Danderyd, Sweden) at a sampling rate of 50 Hz. Two

cameras were used to record participants' foot behaviors (Camera 1) and the overall experimental process (Camera 2), respectively.

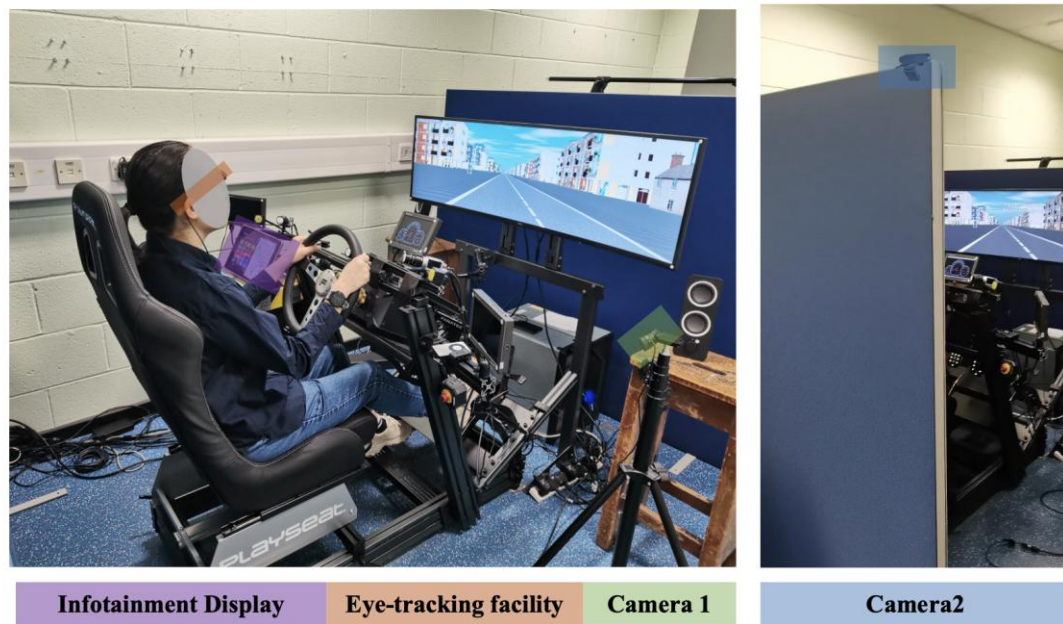


Figure 1. Driving simulator and experimental facilities

2.2.2 Driving scenarios

The driving scenarios were developed using SCANeR Studio 2021 (AVSimulation SAS, France). The simulated scenario consisted of a two-lane urban road divided by a dashed white centerline. Shoulders and pavements were separated from the road by solid white edge lines. To minimize the extra influence of speed differences on drivers' states (Ma et al., 2018), hazard perception performance (Cantwell, 2021), and driving performance (Aarts & Van Schagen, 2006; Favarò et al., 2019), participants were instructed to maintain a target speed of 30 miles per hour (mph) during all drives. This target speed was selected with reference to the UK speed limit in built-up areas (UK Government, n.d.) and is commonly adopted in driving simulator studies (Lee et al., 2024; Lee et al., 2025; Weaver et al., 2023). The simulated environment included buildings, trees, bus stops, streetlamps, pedestrians, oncoming vehicles, and other roadside objects (e.g., signs, parked cars, benches, and construction obstacles).

A hazard was defined as any event or object that causes the driver to consider taking action to prevent a potential danger (Underwood et al., 2013). Each hazard the

participant encountered included a precursor that provided a cue enabling safe drivers to anticipate the upcoming event (Crundall et al., 2012). According to Kahana-Levy et al. (2019), the precursor of each hazard was designed to develop into either a hidden unmaterialized hazard (i.e., an entity was hidden and did not enter the road) or a hidden materialized hazard (i.e., an entity was hidden and then entered the road). Table 2 shows the hazards embedded in the road environment of our simulator. Hazards 1-3 are pedestrian-related while Hazards 4-5 are vehicle-related.

Accordingly, three experimental drives were designed to correspond to the three driving modes. Each drive was 16.5 km long and contained all of the hazards described in Table 2. The onset of each hazard was triggered when the participant's vehicle approached a predetermined distance. To avoid the influence of frequent materialized hazards on drivers' states, there were four hidden unmaterialized hazards and one hidden materialized hazard in each drive. The four hidden unmaterialized hazards were presented in randomized order at varying distances, and their directions (i.e., left or right) were counterbalanced. The hidden materialized hazard consistently appeared near the end of each drive. Moreover, to reduce participants' anticipation of hazards, the same stopped vehicles or construction obstacles associated with the five hazards, but without precursors, were also randomly distributed throughout each scenario. Table A.1 shows detailed information on hazard arrangement, including appearance distance, direction, and sequence for each experimental drive.

Table 2. Hazards and descriptions.

Hazard	Precursor	Hidden unmaterialized	Hidden materialized
Hazard 1	A pedestrian is walking toward the stopped white truck from the pathway between two buildings.	The pedestrian is occluded behind the stopped white truck and remains occluded until the participant's vehicle passes it. 	The pedestrian is initially occluded behind the stopped white truck and then steps into the driver's path. 
Hazard 2	A passenger is walking towards the bus station from the pathway between two buildings.	The pedestrian is occluded behind the stopped bus and remains occluded until the participant's vehicle passes it. 	The pedestrian is initially occluded behind the stopped bus and then steps into the driver's path. 
Hazard 3	A teenager is walking toward the stopped school bus from the pathway between two buildings.	The teenager is occluded behind the stopped school bus and remains occluded until the participant's vehicle passes it. 	The teenager is initially occluded behind the stopped school bus and then steps into the driver's path. 
Hazard 4	A white car is driving toward the stopped red lorry from the pathway between two buildings.	The white car is occluded behind the stopped red lorry and remains occluded until the participant's vehicle passes it. 	The white car is initially occluded behind the stopped red lorry and then enters the driver's path. 
Hazard 5	A red car is driving toward the white construction barrier from the pathway between two buildings.	The red car is occluded behind the white construction barrier and remains occluded until the participant's vehicle passes it. 	The red car is initially occluded behind the white construction barrier and then enters the driver's path. 

2.3 Experimental design

This study employed an experimental design (Berger et al., 2018) to examine the effects of driving mode and pre-drive state on driving safety. Specifically, a 2×3 factorial design was adopted, with driver group and driving mode as the two independent variables. Driver group was treated as a naturally occurring between-subjects factor and included non-nightshift workers, hypothesized to be relatively alert, and nightshift workers, hypothesized to be relatively drowsy. Driving mode was treated as a within-subjects factor and included manual driving, L2 hands-on, and L2 hands-off. Participants in both driver groups completed all three driving mode conditions. This design helped control for inter-individual variability in driver characteristics by allowing comparisons across driving modes within the same participants, thereby increasing internal validity and statistical power (Charness et al., 2012).

2.4 Measures

The effects of driving mode and pre-drive state on driving safety were evaluated from four aspects: drowsiness, attention behaviors, hazard perception performance, and driving performance. The dependent variables and extraction methods were listed in Table 3.

Table 3. Dependent variables and calculations

Measures	Data type	Unit	Data extraction/ calculation
Drowsiness (H1.1, H2.1)			
KSS scores ^P	C	/	Transcribed from paper records into Excel.
Pupil diameter (PD) ^E	C	mm	The average PD during the 60 s before the onset of each hazard.
Attention Behaviors (H1.2, H2.2)			
P _{Infotainment Display} ^E	C	%	$P_{\text{infotainment display}} = \frac{N_{\text{infotainment display}}}{N_{\text{Total}}}$ where $N_{\text{infotainment display}}$ is the number of gaze points located in the infotainment display, and N_{Total} is the total number of gaze points.
P _{Driving Road} ^E	C	%	$P_{\text{Driving road}} = \frac{N_{\text{Driving road}}}{N_{\text{Total}}}$ where $N_{\text{Driving road}}$ is the number of gaze points located in the driving road, N_{Total} is the total number of gaze points.
P _{Instrument Panel} ^E	C	%	$P_{\text{Instrument panel}} = \frac{N_{\text{Instrument panel}}}{N_{\text{Total}}}$ where $N_{\text{instrument panel}}$ is the number of gaze points located in the instrument panel, N_{Total} is the total number of gaze points.
SD of horizontal gaze ^E	C	/	$SD_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$ where $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, x_i is the horizontal gaze coordinate, n is the number of gaze point.
SD of vertical gaze ^E	C	/	$SD_y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}$ where $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, y_i is the vertical gaze coordinate; n is the number of gaze point.
Hazard Perception Performance (H1.3, H2.3)			
Hazard identification ^E	B	/	Within the time window from hazard onset to its offset, the hazard identification was coded as 1 if the driver made at least one fixation (≥ 100 ms) on the hazard AOI, and 0 otherwise.
Time to first fixation on hazard (TTFF) ^E	C	s	$TTFF = t_f - t_0$ where t_0 is the hazard onset time, and t_f is the onset time of the first fixation on the hazard.
Driving Performance (H1.4, H2.4)			
Response ^{DB}	B	/	Within the time window from hazard onset to hazard offset, the response outcome was coded as 1 if the participant responded to the hazard, and 0 otherwise.
TTC _{response} ^{DB}	C	s	$TTC_{\text{response}} = \frac{d(t_r)}{v_{\text{ego}}(t_r) - v_{\text{hazard}}(t_r)}$ Where $d(t_r)$ is distance between the driver's vehicle and the hazard at response time t_r . (this variable was calculated only when a response occurred)
Collision ^{DB}	B	/	Within the time window from the onset of the materialized hazard to its offset, collision was coded as 1 if the participant collided with the hazard, and 0 otherwise.

Note. ^P: paper records; ^E: eye movement data; ^{DB}: driving behavior; C = continuous; B = binary (0 or 1); H = hypothesis.

2.4.1 Drowsiness measures

Driver drowsiness was assessed using the Karolinska Sleepiness Scale (KSS) ratings and pupil diameter (PD). KSS was a 9-point Likert scale ranging from 1 (extremely alert) to 9 (very sleepy, great effort to keep awake, fighting sleep) (Åkerstedt & Gillberg, 1990). It has been used to evaluate drivers' subjective drowsiness before and after a drive. Higher KSS scores (Åkerstedt & Gillberg, 1990) and smaller PD values (Large et al., 2017; Morad et al., 2000; Zhang et al., 2025) indicate higher levels of drowsiness. Previous research has shown that PD size increases following hazard detection relative to the pre-detection period (Baldisserotto et al., 2026). Therefore, to better capture the driver drowsiness while minimizing the influence of hazard-related responses, PD was averaged over the 60 s preceding the onset of each hazard.

2.4.2 Attention behavior measures

Drivers' attention behaviors were characterized in terms of gaze dispersion and gaze spatial distribution. Gaze dispersion was used to compare the spread of gaze across driving modes and driver states. It was measured by the standard deviation (SD) of horizontal gaze and the SD of vertical gaze during each drive. To describe participants' spatial gaze distribution during each drive, the percentage of their gaze points allocated for the infotainment display (i.e., $P_{\text{Infotainment Display}}$), driving road (i.e., $P_{\text{Driving Road}}$), and instrument panel (i.e., $P_{\text{Instrument Panel}}$) were calculated, respectively. It should be noted that these gaze-related indicators were all calculated over the entire drive to align with our research aim of exploring drivers' overall attention behaviors under different driving modes and driver states.

2.4.3 Hazard perception performance measures

Drivers' hazard perception performance was extracted for all hazards in each drive based on eye-movement data, including hazard identification and time to first fixation on hazard (TTFF). Hazard identification referred to whether the driver identified a hazard (fixation duration ≥ 100 ms) within the time window from hazard onset to hazard offset. If a hazard was identified, TTFF was defined as duration from the hazard onset

to the moment when the driver first fixated on the hazard (Crundall et al., 2012). Lower hazard identification probability, or equivalently higher hazard miss probability, and longer TTFF values indicated poorer hazard perception performance (Cao et al., 2022; Crundall et al., 2012).

2.4.4 Driving performance measures

Drivers' driving performance was evaluated using three indicators: response, time to collision at response (TTC_{response}), and collision. First, response referred to whether the driver responded to a hazard between its onset and offset. In the manual driving mode, a response was defined as drivers pressing either the brake or the accelerator pedal. In the two L2 modes, a response was defined as drivers taking over vehicle control by pressing the brake pedal, the accelerator pedal, or the takeover button on the steering wheel. In addition, response outcome was calculated for all hazard in each drive. TTC_{response} referred to the time remained until a collision between the driver's vehicle and the final hazard would occur, assuming that the collision course and speed difference were maintained at the moment the driver responded (Minderhoud & Bovy, 2001). It was calculated only for the last materialized hazard in each drive, and only when a response occurred. Collision referred to whether the driver collided with the materialized hazard, which was calculated for the last materialized hazard in each drive.

2.5 Procedure

The experiment was conducted in a windowless room where the ambient lighting was kept constant for all participants. This arrangement helped minimize the influence of luminance on PD values (Winn & Whitaker, 1994).

The experimental procedure is shown in Figure 2. Upon arrival, all participants read the participant information sheet and sign an informed consent form. They then completed the ESS to assess their general sleepiness tendency and reported the time they had last woken up prior to the experiment. Next, participants completed a 5- to 10-minute practice session to familiarize themselves with the three driving modes. In the manual driving mode, participants had full control of the vehicle and were instructed to maintain a target speed of 30 mph. If they exceeded the speed limit, a “Overspeed Warning” warning appeared at the top of the display until the speed was reduced. In the two L2 driving modes, participants pressed a green button on the left side of the steering wheel to engage automation. Once activated, the system automatically maintained the vehicle’s speed at 30 mph and kept the vehicle centered in the left lane. Participants were instructed to takeover via pressing the first black button on the right side of the steering wheel, pressing the brake pedal, or pressing the accelerator pedal. Textual information indicating the automation status (“Automation On” or “Automation Off”) appeared at the top center of the screen. This information was displayed for 3 s to reduce drivers’ visual distraction. In addition, an automation-status icon remained continuously visible on the instrument cluster to indicate the current state of the system. In the L2 hands-on mode, participants were instructed to keep at least one hand on the upper-middle section of the steering wheel. They were required to remove both hands from the steering wheel except when taking over if they were driving using the hands-off mode. In both L2 driving modes, participants were instructed to keep their feet off the pedals unless taking over.

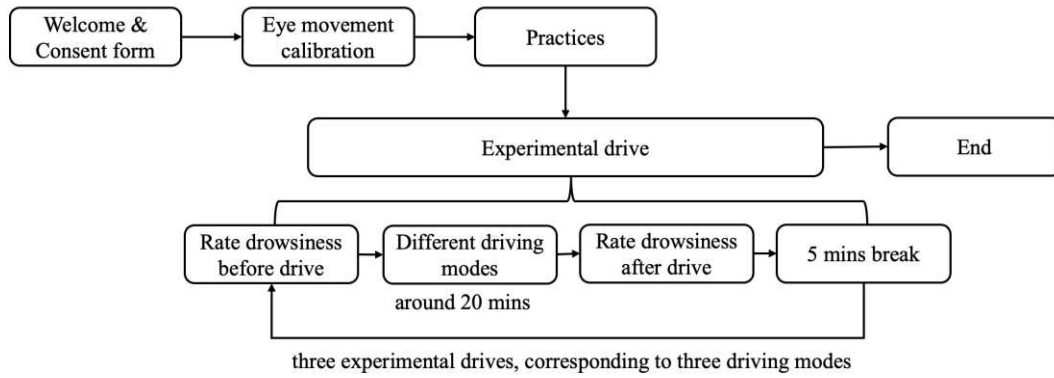


Figure 2. Experiment procedure

During the experimental drive session, participants were instructed to monitor the driving scenario and ensure driving safety. Given that drivers may divert their attention away from the road and engage in NDRTs during L2 driving (Noble et al., 2021), participants were allowed to use the infotainment display during the experiment to better reflect real-world in-vehicle conditions. However, they were also reminded to prioritize driving safety at all times. Each participant completed three experimental drives, corresponding to the three driving modes. The sequence of the driving modes was counterbalanced using a Latin square design. Before each drive, participants rated their current drowsiness level using the KSS. Each drive lasted approximately 20 min and was followed by an after-drive KSS report and a 5-minute break between drives. To reduce the carryover effects of drowsiness from the previous drive, participants could walk around the laboratory during the breaks (as employed by Merat & Jamson, 2013). The entire experiment lasted approximately 2 h. Upon completion, non-nightshift workers received £30 in cash for full participation, whereas nightshift workers received an additional £20 (£10 for the extra effort of participating after a night shift and £10 for transport costs). All the above information was detailed in the participant information sheet.

2.6 Data analysis

2.6.1 Data preprocessing

To enable manipulation checks and drowsiness evaluation, participants' KSS scores before (before-drive KSS scores) and after each drive (after-drive KSS scores) were transcribed into a data sheet and included in the analysis. Eye-movement indicators, including PD, hazard identification, TTFF, SD of horizontal gaze, SD of vertical gaze, $P_{\text{Infotainment Display}}$, $P_{\text{Driving Road}}$ and $P_{\text{Instrument Panel}}$ were extracted using Tobii Pro Lab software (Tobii Inc., Danderyd, Stockholm, Sweden). Driving performance indicators, including response, TTC_{response} , and Collision, were extracted using Python 3.12.0 based on the data output from SCANeR Studio. The extraction methods for all dependent variables were listed in Table 3.

2.6.2 Data analysis methods

Statistical analyses were performed using R Studio (version 2023.10.31). Differences between the two driver groups in demographic variables, including age, driving years, annual mileage, daytime sleepiness tendency, and awake duration, were analyzed using independent-samples t tests.

For continuous dependent variables, linear mixed models (LMMs) including both fixed and random effects were fitted (Baayen et al., 2008; Wan & Sarter, 2022). The normality of model residuals was assessed using Q-Q plots and the one-sample Kolmogorov-Smirnov test (Venables & Ripley, 2002). If the model assumptions were violated, data were log-transformed. If a main effect or interaction was significant, post hoc comparisons were conducted using the “emmeans” package in R (Lenth et al., 2019). The significance level for all analyses was $\alpha = 0.05$, which is a conventional standard in research (Fisher, 1925). Because multiple tests were conducted, Bonferroni correction was applied to control the Type I error rate (Armstrong, 2014; Lehmann & Romano, 2008). The partial eta squared value (η_p^2) was reported as the effect size for LMMs analyses.

For binary dependent variables (i.e., outcome coded as 0 or 1), generalized linear mixed models (GLMMs) with logit link function were employed. In cases of complete data separation, the bias-reduced generalized linear models (GLMs) was used instead

(Kosmidis et al., 2020, 2025). The odds ratios were calculated for GLMMs or GLMs analyses.

2.6.3 Data analysis steps

To verify the manipulation of pre-drive states across driver groups, awake duration and before-drive KSS scores were compared between the non-nightshift and nightshift groups. The follow-up analyses were then conducted in two steps: omnibus tests and separate analyses.

- **Step 1: Omnibus tests**

Omnibus tests were conducted to test H1.1 to H1.4. Models were fitted with driving mode, driver group, and their interaction as fixed effects, and participant ID as a random effect. When a main effect or interaction effect was significant, post hoc analyses were conducted accordingly. These tests showed an overview of effects.

- **Step 2: Separate analyses**

Separate analyses were conducted to test H2.1 to H2.4 and to examine whether the effects of driving mode varied across pre-drive states. Models with driving mode as a fixed effect and participant ID as a random effect were fitted separately for the two driver groups, regardless of whether the interaction between driving mode and driver group was significant. This approach is consistent with recommendations that statistical analyses should align closely with the research questions and a priori hypotheses rather than relying solely on omnibus tests (Hays, 1973; Robinson et al., 2013; Schad et al., 2020). Prior work has also suggested that separate analyses may be more sensitive than omnibus interaction tests for detecting theoretically meaningful group-specific effects (Robinson et al., 2013). If the effects of driving mode on the dependent variables showed a similar directional pattern across the two driver groups, additional planned contrast tests were conducted to examine whether these effects were more pronounced in nightshift workers. Data and analysis code can be found at the following link:

<https://github.com/zhangweibetty/Hand-Engagement-in-Partially-Automated-Driving.git>

3 Results

3.1 Manipulation check

The manipulation for different driver groups was successful in this study. Table 4 shows that nightshift workers had significant longer awake duration than non-nightshift workers ($t(36) = 13.212, p = 0.001$). Moreover, the main effect of driver group on the before-drive KSS scores was significant ($F(1, 36) = 8.786, p = 0.005, \eta_p^2 = 0.900$). Nightshift workers reported higher before-drive KSS scores than non-nightshift workers in the manual driving ($p = 0.018$), L2 hands-on ($p = 0.002$), and L2 hands-off ($p = 0.045$) modes. In addition, before-drive KSS scores did not differ across the three driving modes before the drives in either driver group, indicating that drivers within each group remained in a relatively consistent state before each drive.

Table 4. Awake duration and before-drive KSS scores for the two driver groups.

	Awake duration (s)	Before-drive KSS scores		
		Manual	L2 hands-on	L2 hands-off
Non-nightshift workers	6.956 (0.309)	2.947(0.393)	2.789 (0.271)	3.263 (0.304)
Nightshift workers	18.551 (0.815)	4.632 (0.553)	4.526 (0.455)	4.526 (0.526)

Note. Values are means with standard errors in parentheses.

3.2 Overview of data reporting

Mean (M) and standard error (SE) of each dependent variable are presented in Table 5. Standard error were reported to reflect the precision of the estimated means and to facilitate comparisons across conditions (Lee et al., 2015; Barde & Barde, 2012), whereas inferential conclusions were based on the LMM and GLMM results.

Results are reported from four aspects: driver drowsiness, attention behaviors, hazard perception performance, and driving performance. For each aspect, the omnibus test results (Step 1) are presented first, with the corresponding results for continuous and binary dependent variables shown in Tables A.2 and A.3, respectively. The separate analyses results (Step 2) are then presented, with the corresponding results shown in Tables 6 and 7, respectively.

Table 5. Mean and standard errors of all dependent variables for both driver groups.

	Non-nightshift workers			Nightshift workers		
	Manual	L2 hands-on	L2 hands-off	Manual	L2 hands-on	L2 hands-off
Drowsiness						
After-drive KSS scores	3.789 (0.409)	4.421 (0.435)	4.684 (0.502)	5.474 (0.492)	5.763 (0.469)	6.184 (0.542)
PD (mm)	3.145 (0.092)	3.142 (0.086)	3.096 (0.083)	3.255 (0.103)	3.233 (0.112)	3.233 (0.047)
Attention behaviors						
P _{Infotainment Display} (%)	0.404 (0.129)	0.848 (0.370)	0.632 (0.166)	0.645 (0.206)	0.954 (0.259)	0.991 (0.305)
P _{Driving Road} (%)	88.704 (2.176)	94.916 (1.663)	96.037 (0.906)	91.145 (1.912)	93.894 (1.148)	94.780 (1.167)
P _{Instrument Panel} (%)	4.994 (1.747)	0.675 (0.112)	0.569 (0.112)	4.017 (1.437)	0.736 (0.151)	0.719 (0.219)
SD of horizontal gaze	97.944 (4.674)	118.002 (5.171)	123.057 (5.176)	99.982 (5.141)	116.424 (6.059)	120.383 (8.032)
SD of vertical gaze	124.822 (4.731)	92.816 (5.150)	91.685 (5.442)	133.551 (4.769)	104.468 (6.009)	100.707 (6.781)
Hazard perception performance						
Hazar miss probability (%)	0.000 (0.000)	2.174 (1.520)	3.261 (1.852)	8.696 (2.938)	6.522 (2.574)	11.364 (3.383)
TTFF (s)	3.076 (0.277)	2.996 (0.292)	2.492 (0.230)	4.553 (0.326)	4.495 (0.337)	4.462 (0.408)
Driving performance						
Response probability (%)	29.474 (4.678)	34.736 (4.885)	42.105 (5.066)	36.842 (4.949)	30.526 (4.725)	36.842 (4.949)
TTC _{response} (s)	4.805 (0.858)	6.627 (0.949)	7.603 (1.078)	7.332 (1.279)	7.165 (1.027)	7.768 (1.340)
Collision probability (%)	0.000 (0.000)	15.789 (8.365)	10.526 (7.041)	5.263 (5.123)	31.578 (10.664)	36.842 (11.066)

Note. Values are Mean (Standard Error); PD: pupil diameter; TTFF: time to first fixation on hazard; P_{Infotainment Display}: percentage of gaze points allocated to infotainment display. P_{Driving Road}: percentage of gaze points allocated to driving road; P_{Instrument Panel}: percentage of gaze points allocated to instrument panel.

Table 6. Separate analyses (Step 2): main effects of driving mode on continuous dependent variables for both driver groups

Drowsiness indicators	After-drive KSS scores					PD				
	df ₁	df ₂	F	<i>p</i>	η_p^2	df ₁	df ₂	F	<i>p</i>	η_p^2
Main effect of driving mode										
Non-nightshift drivers	2	36	2.974	0.023	0.140	2	256.0	3.658	0.027	0.030
Nightshift drivers	2	36	2.545	0.031	0.120	1	260	1.882	0.154	0.010
Hazard perception performance	TTF									
	df ₁	df ₂	F	<i>p</i>	η_p^2					
Main effect of driving mode										
Non-nightshift drivers	2	248.98	1.464	0.233	0.010					
Nightshift drivers	2	231.52	0.017	0.983	0.000					
Gaze behaviors	P _{Infotainment Display}					P _{Driving Road}				
	df ₁	df ₂	F	<i>p</i>	η_p^2	df ₁	df ₂	F	<i>p</i>	η_p^2
Main effect of driving mode										
Non-nightshift drivers	2	36	3.273	0.049	0.130	2	36	7.521	0.002	0.290
Nightshift drivers	2	36	4.674	0.016	0.180	2	36	2.439	0.102	0.120
	P _{Instrument Panel}									
	df ₁	df ₂	F	<i>p</i>	η_p^2					
Main effect of driving mode										
Non-nightshift drivers	2	36	5.889	0.006	0.250					
Nightshift drivers	2	36	3.659	0.036	0.170					
	SD of horizontal gaze					SD of vertical gaze				
	df ₁	df ₂	F	<i>p</i>	η_p^2	df ₁	df ₂	F	<i>p</i>	η_p^2
Main effect of driving mode										
Non-nightshift drivers	2	36	15.073	<.001	0.460	2	36	19.212	<.001	0.520
Nightshift drivers	2	36	8.037	0.001	0.310	2	36	26.188	<.001	0.590
Driving performance	TTC _{response}									
	df ₁	df ₂	F	<i>p</i>	η_p^2					
Main effect of driving mode										
Non-nightshift drivers	2	33.70	6.667	0.004	0.280					
Nightshift drivers	2	25.42	0.052	0.950	0.004					

Note. η_p^2 indicates the partial eta square, determining the effect size. Degrees of freedom for fixed effects were approximated using Satterthwaite's method, which can yield non-integer denominator degrees of freedom.

Table 7. Separate analyses (Step 2): main effects of driving mode on binary dependent variables for both driver groups

Predictors	Coefficient (β)	SE	Z	p	OR	CI-L	CI-U
Hazard identification							
<i>Non-nightshift Group</i>							
Intercept	3.241	0.548	5.917	< .001	25.571	8.738	74.830
Driving Mode [L2 hands-on]	0.348	0.846	0.411	0.681	1.416	0.270	7.431
Driving Mode [Manual]	1.946	1.528	1.274	0.203	7.000	0.351	139.800
<i>Nightshift Group</i>							
Intercept	2.290	0.441	5.195	< .001	9.876	4.163	23.432
Driving Mode [L2 hands-on]	0.649	0.554	1.172	0.241	1.914	0.646	5.670
Driving Mode [Manual]	0.318	0.515	0.617	0.537	1.374	0.501	3.768
Response							
<i>Non-nightshift Group</i>							
Intercept	-0.341	0.322	-1.058	0.290	0.711	0.379	1.337
Driving Mode [L2 hands-on]	-0.382	0.329	-1.161	0.246	0.683	0.358	1.300
Driving Mode [Manual]	-0.678	0.338	-2.009	0.045	0.508	0.262	0.984
<i>Nightshift Group</i>							
Intercept	-0.653	0.373	-1.752	0.080	0.521	0.251	1.080
Driving Mode [L2 hands-on]	-0.368	0.347	-1.061	0.289	0.692	0.351	1.366
Driving Mode [Manual]	0.000	0.338	0.000	1.000	1.000	0.515	1.941
Collision							
<i>Non-nightshift Group</i>							
Intercept	-1.946	0.694	-2.805	0.005	0.143	0.037	0.556
Driving Mode [L2 hands-on]	0.395	0.920	0.430	0.667	1.485	0.245	9.005
Driving Mode [Manual]	-1.718	1.625	-1.057	0.290	0.179	0.007	4.337
<i>Nightshift Group</i>							
Intercept	-0.511	0.474	-1.078	0.281	0.600	0.237	1.519
Driving Mode [L2 hands-on]	-0.220	0.682	-0.323	0.747	0.802	0.211	3.052
Driving Mode [Manual]*	-2.001	0.992	-2.018	0.044	0.135	0.019	0.944

Note. Probability distribution: Binomial (Logit). Hazard identified = 1; Not identified = 0; Responded = 1; Did not respond = 0; Collided = 1; Did not collide = 0.

3.3 Drowsiness

3.3.1 KSS scores

Omnibus tests. As shown in the Table A.2, the main effect of driver group on after-drive KSS scores was significant. Nightshift workers ($M = 5.184$, $SE = 0.519$) reported higher after-drive KSS scores than non-nightshift workers ($M = 3.649$, $SE = 0.418$) ($p = 0.001$). There was a significant main effect of driving mode on after-drive KSS scores. The post hoc analysis revealed higher after-drive KSS scores in L2 hands-off mode than in manual driving mode ($p = 0.006$), indicating that L2 hands-off induced greater drowsiness.

Separate analyses. Figure 3 shows that non-nightshift workers exhibited more consistent after-drive KSS scores (shorter violins) than nightshift workers. The main effect of driving mode on after-drive KSS scores was significant for both groups (see Table 6). Non-nightshift workers reported significantly higher after-drive KSS scores in the L2 hands-off mode than in the manual driving mode ($p = 0.023$). Nightshift workers also reported significantly higher after-drive KSS scores in the L2 hands-off mode than the manual driving mode ($p = 0.031$). Notably, data points for nightshift workers in the L2 hands-off mode were more concentrated at higher after-drive KSS scores than participants in the other group. Planned contrast analysis further showed that the difference in after-drive KSS scores between the hands-off and manual driving modes did not significantly differ between the participants in the two groups ($t(72) = -0.427$, $p = 0.670$). This indicates that the increase in drowsiness in the hands-off mode relative to manual driving was not more pronounced in nightshift workers.

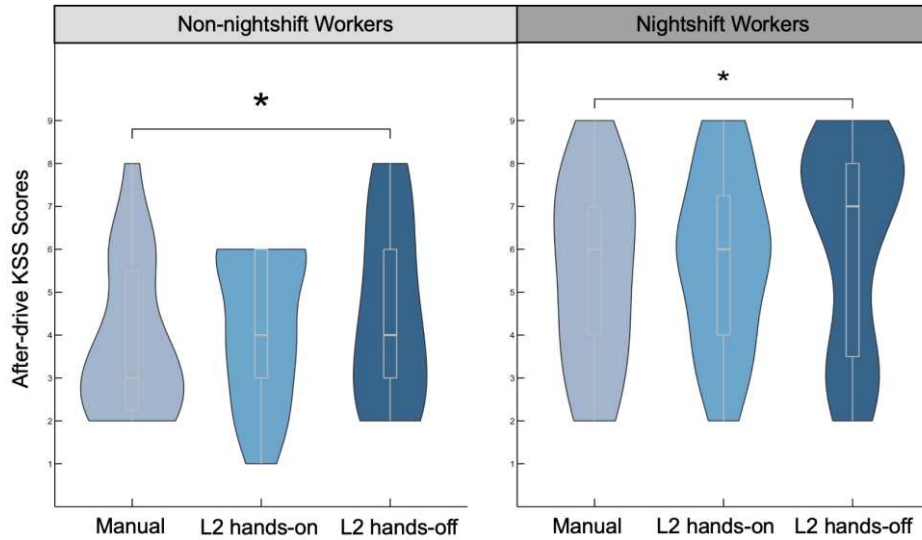


Figure 3. Violin plots of after-drive KSS scores for different driving modes.

Additionally, this study examined whether drivers' drowsiness increased over time by comparing KSS scores obtained before and after each drive. The results showed that after-drive KSS scores ($M = 5.053$, $SE = 0.503$) were significantly higher than before-drive KSS scores ($M = 3.780$, $SE = 0.459$) ($F_{(1, 182)} = 66.387$, $p < 0.001$, $\eta_p^2 = 0.270$). A similar pattern, with higher after-drive than before-drive KSS scores, was observed for non-nightshift workers in the L2 hands-off ($p = 0.012$) and L2 hands-on ($p = 0.004$) modes, and for nightshift workers in the L2 hands-off ($p = 0.023$) modes.

3.3.2 Pupil Diameter (PD)

Omnibus tests. As shown in Table A.2, results suggested a main effect of driving mode on PD. The L2 hands-off mode ($M = 3.166$, $SE = 0.068$) resulted in a smaller PD than both manual driving ($M = 3.201$, $SE = 0.070$) ($p = 0.021$) and L2 hands-on ($M = 3.188$, $SE = 0.071$) ($p = 0.032$).

Separate analyses. When analyzed by driver group, the main effect of driving mode on PD was significant for non-nightshift workers (see Table 6). As shown in Figure 4, PD in the L2 hands-off mode was smaller than in the L2 hands-on ($p = 0.024$) and manual driving modes ($p = 0.017$), indicating greater sleepiness in the L2 hands-off mode. For nightshift workers, the main effect of driving mode on PD was not significant.

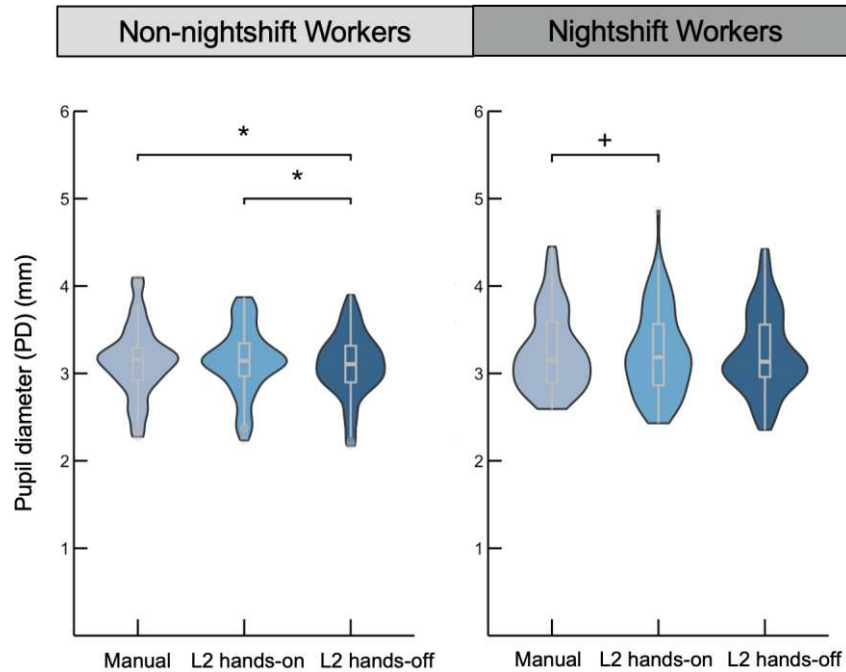


Figure 4. Violin plots of PD for different driving modes.

3.4 Attention behaviors

3.4.1 Gaze dispersion

(1) SD of horizontal gaze

Omnibus tests. For SD of horizontal gaze, the main effect of driving mode was significant (see Table A.2). Compared with manual driving ($M = 98.963$, $SE = 4.852$), both the L2 hands-on mode ($M = 117.213$, $SE = 5.559$, $p = 0.001$) and the L2 hands-off mode ($M = 121.720$, $SE = 6.672$, $p = 0.001$) produced greater horizontal gaze dispersion.

Separate analyses. When analyzed by driver group, the main effect of driving mode on SD of horizontal gaze was significant for participants in both groups (see Table 6). As shown in Figure 5 a), a similar overall pattern of driving mode was observed in the two groups, with both the hands-on and -off modes leading to greater horizontal gaze dispersion than the manual driving mode. Planned contrast analysis further indicated that the hands-off versus manual driving effect on SD of horizontal gaze was significantly larger in non-nightshift workers than in nightshift workers ($t(72) = -2.221$, $p = 0.030$).

(2) SD of vertical gaze

Omnibus tests. For SD of vertical gaze, the main effect of driving mode was also significant (see Table A.2). Compared with manual driving ($M = 129.187$, $SE = 4.794$), both the L2 hands-on mode ($M = 98.642$, $SE = 5.684$, $p < 0.001$) and the hands-off mode ($M = 96.196$, $SE = 6.155$, $p < 0.001$) produced smaller vertical gaze dispersion.

Separate analyses. When analyzed by driver group, the main effect of driving mode on SD of vertical gaze was significant for both participant groups (see Table 6). As shown in Figure 5 b), a similar pattern of driving mode was observed in the two groups, with both the hands-on and -off modes leading to smaller vertical gaze dispersion than manual driving. Planned contrast analysis further indicated that the hands-off versus manual driving effect on vertical gaze dispersion was significantly larger in non-nightshift workers than workers in the other group ($t(72) = 3.601$, $p = 0.001$).

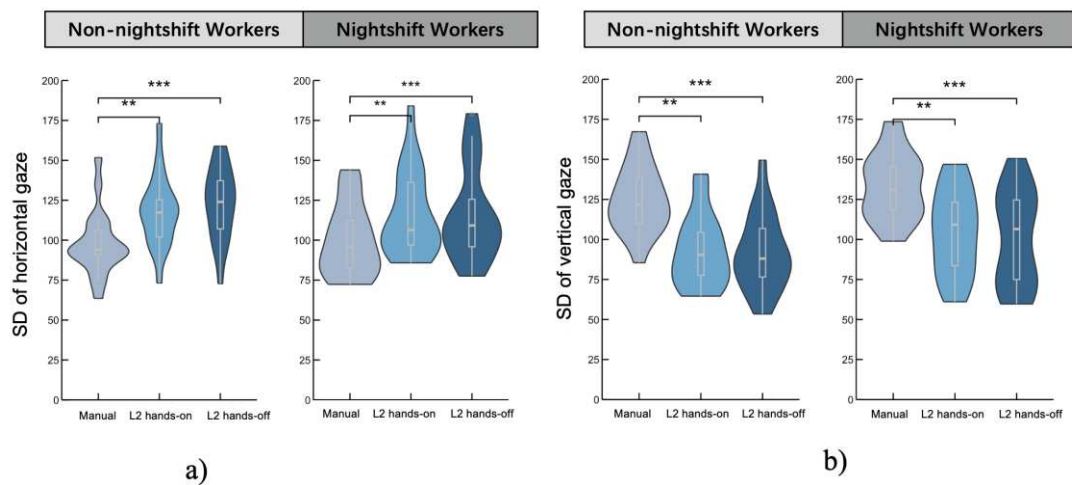


Figure 5. Violin plots of SD of horizontal and vertical gaze.

3.4.2 Gaze spatial distribution

Participants' gaze point distributions are shown in Figure 6.

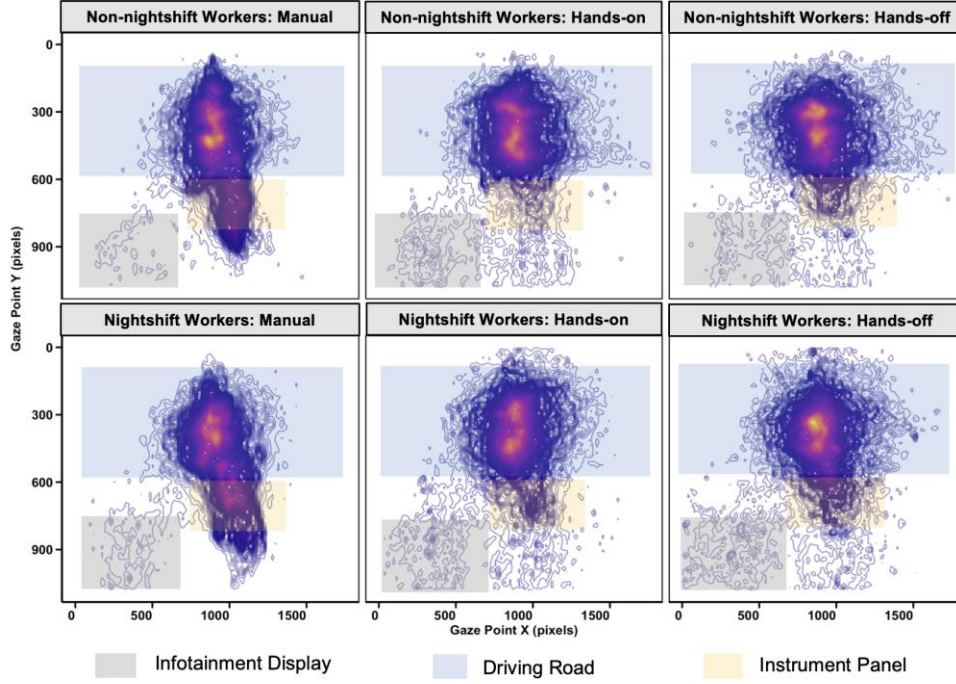


Figure 6. Contour plots of participants' gaze points.

(1) $P_{\text{Infotainment Display}}$

Omnibus tests. The results indicated significant main effects of driving mode on $P_{\text{Infotainment Display}}$ (see Table A.2). For infotainment display, both L2 hands-on ($M = 0.901$, $SE = 0.223$, $p = 0.002$) and -off ($M = 0.812$, $SE = 0.174$, $p = 0.035$) resulted in significant higher $P_{\text{Infotainment Display}}$ than manual driving ($M = 0.252$, $SE = 0.122$).

Separate analyses. Results of separate analyses by driver group (see Table 6) showed that the main effect of driving mode on $P_{\text{Infotainment Display}}$ was significant for both driver groups. L2 hands-on ($p = 0.050$) and -off mode ($p = 0.023$) produced significant higher $P_{\text{Infotainment Display}}$ than manual driving for non-nightshift workers. Similarly, L2 hands-on ($p = 0.006$) and -off mode ($p = 0.035$) led to significant higher $P_{\text{Infotainment Display}}$ than manual driving in nightshift workers. Further planned contrast analysis indicated that the hands-off versus manual driving effect ($t(72) = -1.857$, $p = 0.067$) and hands-on versus manual driving effect on $P_{\text{Infotainment Display}}$ ($t(72) = 0.308$, $p = 0.759$) was not significantly different.

(2) P Driving Road

Omnibus tests. The results indicated significant main effects of driving mode on P Driving Road (see Table A.2). P Driving Road was significantly higher in L2 hands-on ($M = 94.405$, $SE = 1.000$) ($p = 0.001$) and -off modes ($M = 95.409$, $SE = 0.736$) ($p < 0.001$) compared with manual driving ($M = 89.924$, $SE = 1.443$).

Separate analyses. Significant effects of driving mode on P Driving Road were observed for non-nightshift workers, where both L2 hands-on ($p = 0.005$) and -off modes ($p < 0.001$) resulted in significantly higher gaze percentages than that of manual driving (see Table 6). Further planned contrast analysis indicated that the hands-off versus manual driving effect ($t(72) = -1.977$, $p = 0.052$) and hands-on versus manual driving effect on P Driving Road ($t(72) = -0.128$, $p = 0.898$) was not significantly different.

(3) P Instrument Panel

Omnibus tests. The results indicated significant main effects of driving mode on P Instrument Panel (see Table A.2). Specifically, P Instrument Panel was significantly higher in manual driving ($M = 4.505$, $SE = 1.119$) than in L2 hands-on ($M = 0.306$, $SE = 0.093$, $p = 0.006$) and -off modes ($M = 0.644$, $SE = 0.122$, $p < 0.001$).

Separate analyses. Manual driving produced significantly higher P Instrument Panel than L2 hands-on ($p = 0.027$) and -off modes ($p = 0.002$) in non-nightshift workers. Also, P Instrument Panel of manual driving was also significantly higher than L2 hands-off mode for nightshift workers ($p = 0.011$). Planned contrast analysis result indicated that the hands-off versus manual driving effect ($t(72) = 0.823$, $p = 0.413$) and hands-on versus manual driving effect on P Instrument Panel ($t(72) = -0.345$, $p = 0.732$) was not significantly different.

3.5 Hazard perception performance

3.5.1 Hazard identification

Omnibus tests. Results for hazard identification are shown in Table A.3. Nightshift workers showed a significantly higher hazard miss probability than non-nightshift workers (8.824% vs. 1.832%, $OR = 0.292$, $p = 0.033$), indicating poorer hazard

perception performance. The main effect of driving mode on hazard identification was not significant. Nevertheless, descriptively, the hazard miss probability tended to be higher in the L2 hands-off mode than in the other two driving modes.

Separate analyses. Table 7 shows that no significant differences in hazard identification were found among driving modes within either driver group. However, descriptively, workers in both groups showed a higher trend of hazard miss probability in the L2 hands-off mode than in the manual or hands-on modes (see Figure 7).

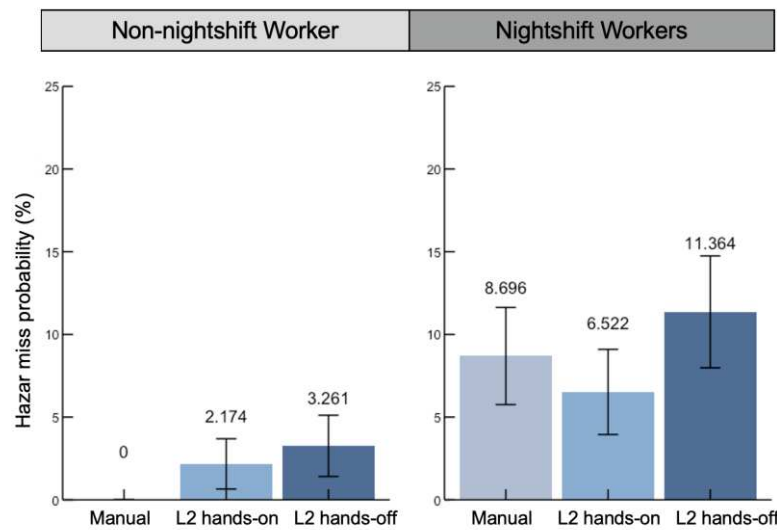


Figure 7. Bar plots of hazard miss probability for different driving modes.

3.5.2 Time to first fixation on hazard (TTFF)

TTFF values were extracted only for hazards that were successfully identified. The complete number of TTFF observations in each condition was supposed to be 95. For non-nightshift workers, the numbers of valid TTFF observations in the manual driving, L2 hands-on, and L2 hands-off conditions were 95, 93, and 92, respectively. For nightshift workers, the corresponding numbers were 87, 89, and 85.

Omnibus tests. As shown in the Table A.2, the main effect of driver group on TTFF was significant. Nightshift workers ($M = 4.504$, $SE = 0.740$) had significantly higher TTFF than non-nightshift workers ($M = 2.855$, $SE = 0.582$) ($p = 0.001$). Figure 8 shows that data for non-nightshift and nightshift workers were concentrated at lower and higher TTFF values, respectively. This implied that nightshift drivers needed more time to identify a potential hazard.

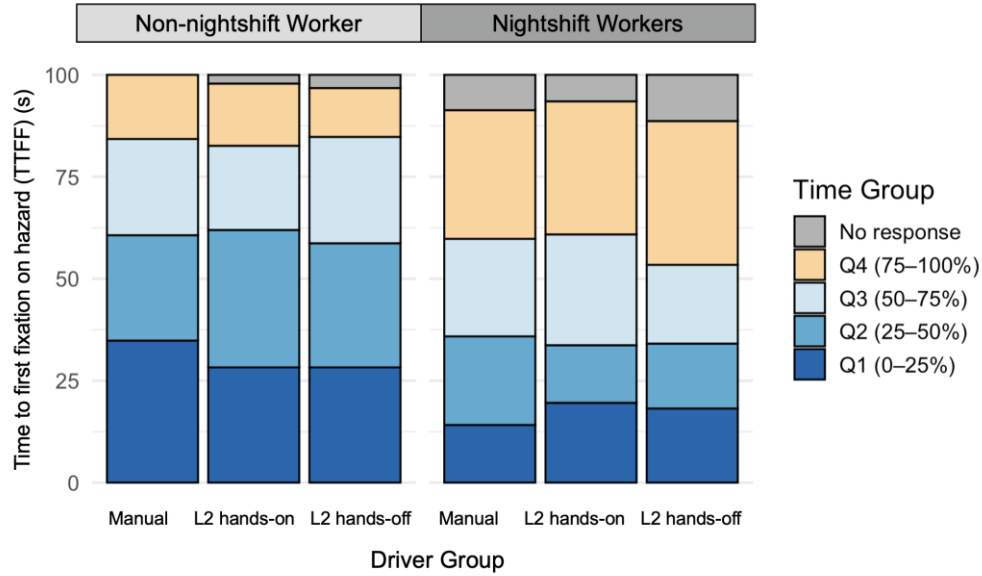


Figure 8. Stacked bar plots of quantiles of time to first fixation on hazard (quantiles were calculated by driving mode).

Separate analyses. The main effect of driving mode on TTFF was not significant for workers in either group (see Table 6). However, the hands-off mode for nightshift workers tended to be associated with longer fixation times than hands-on and manual driving modes (see Figure 8).

3.6 Driving performance

3.6.1 Response

Omnibus tests. Results for response are shown in Table A.3. Drivers were significantly more likely to respond in the L2 hands-off mode than in the manual driving mode (39.474% vs. 33.158%, OR = 0.501, $p = 0.042$).

Separate analyses. When analyzed separately, the higher response probability in the L2 hands-off mode than in the manual driving mode was observed only in non-nightshift workers (see Table 7). Figure 9 shows that non-nightshift workers showed a higher response probability in the L2 hands-off mode than in the manual driving mode ($p = 0.045$). No significant differences in response were found among the manual, L2 hands-on, and L2 hands-off modes for nightshift workers.

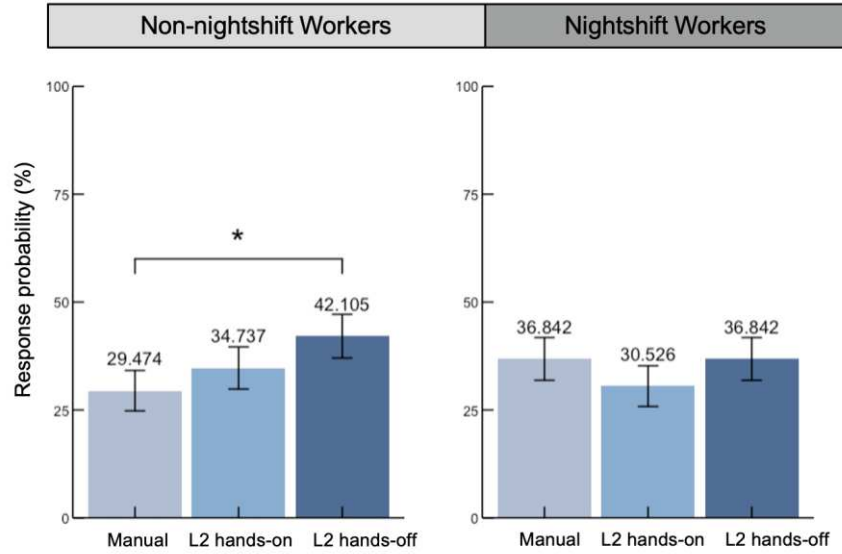


Figure 9. Bar plots of response probability for different driving modes.

3.6.2 TTC_{response}

TTC_{response} values were extracted only when a response occurred. The complete number of TTC_{response} observations in each condition was supposed to be 95. For non-nightshift workers, the numbers of valid observations in the manual driving, L2 hands-on, and -off conditions were 26, 33, and 39, respectively. The corresponding numbers for nightshift workers were 33, 27, and 35, respectively.

Omnibus tests. Results of driver group, driving mode, and their interaction effects were listed in the Table A.2.

Separate analyses. When analyzed by driver group, the main effect of driving mode on TTC_{response} was significant for non-nightshift workers but not for the other group (see Table 6). Figure 10 shows that non-nightshift workers exhibited significantly higher TTC_{response} in the L2 hands-on and -off modes than in manual driving. This implies safer responses in the two L2 driving modes.

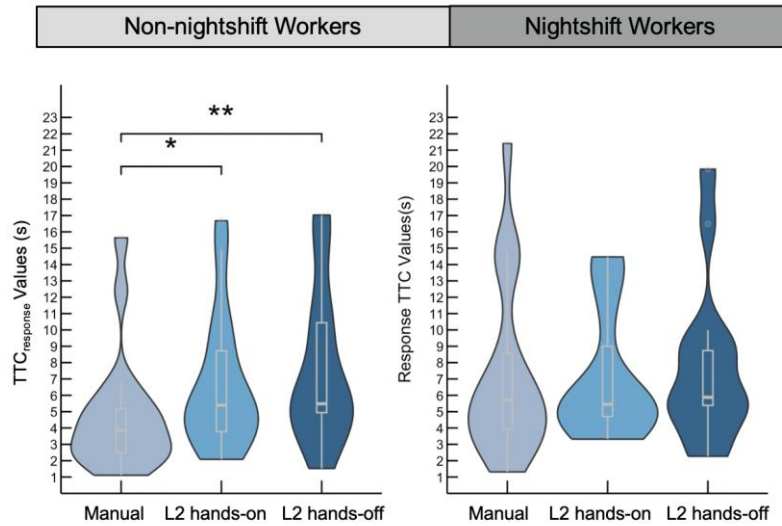


Figure 10. Violin plots of TTC_{response} in different driving modes.

3.6.3 Collision

Omnibus tests. Table A.3 shows the effects of driver group, driving mode, and their interaction effects on collision.

Separate analyses. No significant differences in collision were found among the three driving modes for non-nightshift workers (see Table 7). However, for nightshift workers, L2 hands-off mode induced significantly higher collision probability than in the manual driving mode ($p = 0.044$) (see Figure 11).

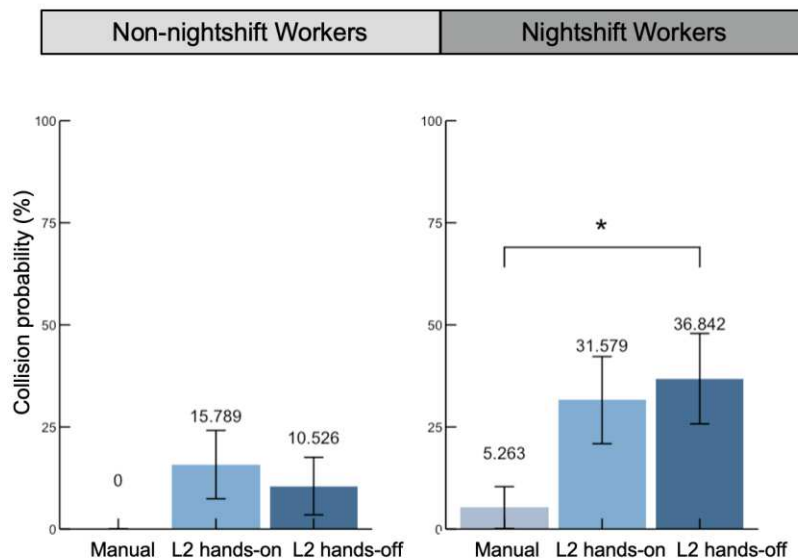


Figure 11. Bar plots of collision probability for different driving modes.

4 Discussion

This study explored how driving mode influenced driver drowsiness, attention behaviors, hazard perception performance, and driving performance considering drivers' pre-drive states. The pre-drive states of non-nightshift and nightshift workers were assumed to be alert and drowsy before driving, respectively. Each participant completed three drives corresponding to manual, L2 hands-on, and L2 hands-off modes. The manipulation check results showed that nightshift workers had longer awake durations and higher pre-drive KSS ratings than non-nightshift workers. This indicates that the manipulation of driver state across the two groups was successful.

4.1 Influence of driving mode

Some of the first set of hypotheses (H1) was supported. This proposed that, among the three driving modes, hands-off driving would be associated with the poorest driving safety outcomes in terms of driver drowsiness, attention behaviors, hazard perception performance, and driving performance. Regarding drowsiness, this study found that drivers' drowsiness increased with time, where after-drive KSS scores were higher than before-drive KSS scores. Moreover, the hands-off mode led to a drowsier state than manual driving. Specifically, after-drive KSS scores were significantly higher, and PD values were significantly lower in the hands-off mode than in the manual driving mode. Therefore, H1.1 was supported, suggesting that reduced hand engagement in L2 driving may contribute to increased driver drowsiness. This finding is consistent with previous studies showing that autonomous driving generally induces greater drowsiness than manual driving (Pan, Logan, et al., 2025; Vogelpohl et al., 2019).

The L2 hands-on and -off modes showed similar patterns of attention behavior. Compared with manual driving, both L2 modes were associated with greater horizontal gaze dispersion and smaller vertical gaze dispersion. This pattern is reasonable because drivers still needed to actively control lane position (i.e., smaller horizontal gaze dispersion) and maintain speed (i.e., greater vertical gaze dispersion) during manual driving, whereas these demands were reduced in L2 driving. This finding is consistent

with previous research reporting greater horizontal gaze dispersion in highly automated driving than in manual driving (Damböck et al., 2013). In terms of gaze allocation, the spatial distribution results suggested that both L2 driving modes led drivers to devote more attention to the infotainment display than manual driving. These findings are consistent with those of Lee et al. (2024), who reported that longer hands-off durations were associated with reduced on-road glance times. They are also aligned with the results in the previous studies that drivers are more likely to engage in NDRTs during L2 driving (Carsten et al., 2012; Gershon et al., 2023). Therefore, H1.2 was supported.

Although driving mode showed no statistically significant effect on hazard perception performance, a declining trend in hazard perception performance was observed from manual driving to the L2 hands-on and hands-off modes. That is, compared with the manual driving mode, the results suggested a trend toward a higher probability of failing to identify hazard in both L2 modes. This finding was partly aligned with the findings of Zangi et al. (2022), which found that drivers were less aware of road hazards when their attention was diverted from the road. In addition, the non-significant result might be attributed to the small number of hazards in each drive to reach statistical significance. Therefore, although H1.3 was not statistically supported, the observed patterns were in the expected direction.

Finally, although the effect of driving mode was not significant for most driving performance outcomes, the overall trend in related dependent variables suggested reduced driving safety from manual driving to L2 hands-on and -off modes. Specifically, the L2 hands-off mode showed a trend of higher collision probability than the hands-on and manual driving modes. Therefore, although H1.4 was not statistically supported, the observed patterns were again in the expected direction.

4.2 The influence of driving mode across pre-drive states

Considering drivers' pre-drive states, the second set of hypotheses (H2) concerning the effects of driving modes on driving safety was not supported, except for the hypothesis related to driving performance. Regarding driver drowsiness, the separate

analyses of after-drive KSS scores suggested that the L2 hands-off mode led to a drowsier state than manual driving for both groups. However, the contrast analysis results showed that the adverse effect of hands-off relative to manual driving did not differ significantly between the two driver groups. These findings did not support H2.1, which proposed that the adverse effect of the L2 hands-off mode would be more pronounced for drivers in a pre-drive drowsy state. These results also contrast with those of Vogelpohl et al. (2019), who found that the increase in drowsiness was greater in automated driving than in manual driving and was especially pronounced under sleep-restricted conditions. One possible explanation is that the level of drowsiness in the drowsy group of Vogelpohl et al. (2019) may have been greater than that of the nightshift workers in the current study, making their drowsiness more likely to worsen over time. That is, the pre-drive drowsy participants in their study were restricted to a maximum of 5 h of sleep before the experiment, whereas the nightshift workers in the present study came to the laboratory immediately after completing a night shift. It is possible that the nightshift workers in our study had adapted to their shift schedules and were therefore not as drowsy as the drivers in Vogelpohl et al. (2019).

In addition, keeping hands on the steering wheel may help mitigate drowsiness, but only for non-nightshift workers who were initially alert before driving. Specifically, a trend of relatively lower after-drive KSS scores and higher PD values was observed in the hands-on mode than in the hands-off mode for non-nightshift workers. This pattern was not found among nightshift workers. When the hands-off mode induced drowsiness through task underload, placing hands on the steering wheel might help non-nightshift workers remain more engaged in the driving task. However, when drivers were already in a drowsier state, the hands-on mode appeared to offer little benefit in counteracting drowsiness among nightshift workers.

Regarding attention behaviors, the greater diversion of attention away from the road in the L2 hands-off mode was not observed in drowsy drivers (i.e., H2.2 was not supported). That is, participants showed greater horizontal gaze dispersion and fewer

vertical gaze dispersion in two L2 modes than in the manual driving modes, and this effect was more pronounced for non-nightshift workers. Moreover, participants allocated more attention to the infotainment display in the two L2 modes than in the manual driving mode, while this effect did not differ between nightshift workers and non-nightshift workers. These findings suggested that drivers in a pre-drive drowsy state did not show a stronger tendency to allocate attention to NDRTs than those in an alert state. This was somewhat unexpected, given previous findings suggesting that engaging in NDRTs may help maintain alertness during prolonged automated driving (Vogelpohl et al., 2019; Pan, Logan, et al., 2025). One possible explanation is that the level of drowsiness among the nightshift workers in the present study may not have been sufficiently high to induce greater engagement with NDRTs as a compensatory strategy.

Regarding hazard perception performance, no significant differences in the effects of driving mode between the two driver groups were found (i.e., H2.3 was not supported). This finding was also inconsistent with Watling and Home (2022), who reported that hazard perception performance declined with driving time, with the effect being more pronounced under sleep-deprived conditions. One possible explanation is that the drive durations in the present study (i.e., 20 minutes) were shorter than those in Watling and Home (2022) (i.e., 60 minutes), and therefore the effect of pre-drive drowsiness may not have fully emerged.

Finally, poorer driving performance in the L2 hands-off mode relative to manual driving was more evident in nightshift workers than in the other group (i.e., H2.4 was supported). More specifically, non-nightshift workers responded more frequently to hazards and showed higher TTC_{response} in the hands-off mode than in manual driving. As a result, their collision rate in the hands-off mode was reduced. In contrast, nightshift workers showed similar response frequencies across driving modes, which resulted in a significantly higher collision rate in the hands-off mode. These findings suggest that non-nightshift workers adopted a more cautious response strategy than their nightshift

counterparts when driving in the hands-off mode. This might be because nightshift workers were generally drowsier than non-nightshift workers. Drowsiness-related cognitive impairment (Banks & Dinges, 2007) may have reduced drivers' ability to accurately judge the automation's capability and the current driving situation, thereby amplifying the detrimental effect of the hands-off mode on driving performance.

4.3 Implications and Limitations

With the deployment of L2 hands-off systems, the present study may have several implications for regulating the use of the hands-off mode. Given the adverse effects of the hands-off mode on driver drowsiness and attention behaviors, as well as the detrimental trends in hazard perception performance and driving performance, restrictions on its use may be necessary. For example, the hands-off mode could be deactivated under drowsiness-inducing conditions (e.g., motorways and night driving) (May & Baldwin, 2009) and in situations with frequent hazards (e.g., crowded urban areas and peak traffic hours). In addition, because alert drivers may benefit from placing their hands on the steering wheel to maintain alertness, drivers should be prompted to resume hand contact with the steering wheel after using the hands-off mode for a limited period of time. Moreover, nightshift workers seemed to adopt effective strategies less in the hands-off mode to avoid collisions. Assessing drivers' pre-drive drowsiness before allowing the use of the hands-off mode may be an important precaution.

This study has limitations. First, the driving duration was relatively short, with each trial lasted for approximately 20 minutes. Even this period could lead to moderate or high drowsiness for automated driving (Vogelpohl et al., 2019), the influence of hand engagement during longer driving periods should be examined to better understand the development of drowsiness across different driving modes. Second, this study included only few typical hazards. Future studies should consider incorporating a wider range of hazard types, such as divided-attention and focused-attention hazards (Crundall et al., 2012) , to provide a more comprehensive assessment. Third, some elements of the

driving scenarios, such as the shoulder, pavement, side streets, and merging roads were not designed with sufficient realism or clarity to be easily distinguished. For example, the boundary between the shoulder and roadway was not clearly defined when only a solid white line was used. Roadway characteristics such as shoulder width, guardrail presence, and roadway geometry may influence drivers' hazard perception performance and behaviors (Ben-Bassat & Shinar, 2011; Van Der Horst & De Ridder, 2007). Future studies should incorporate more realistic driving scenarios to improve ecological validity.

Fourth, to simulate real-world conditions, participants were allowed to engage in NDRTs while still being instructed to ensure safety. However, engaging in NDRTs may influence driver drowsiness (Pan, Logan, et al., 2025), attention behaviors (Pan, Payre, et al., 2025), hazard perception performance (Zangi et al., 2022), and driving performance (Zhang et al., 2023). Future studies should therefore examine the effects of driving mode and pre-drive state on driving safety under conditions with and without NDRTs.

Lastly, the driving speed in this study was fixed at 30 mph across driver groups and driving modes. While this design helped minimize the confounding influence of speed, it may restrict the generalizability of the present findings to other speed levels. Future research is needed to investigate how the effects of driving mode on driving safety may vary under different speed conditions.

5 Conclusion

With the emerging availability of L2 hands-off mode, it is increasingly important to evaluate its influence on driving safety. The present study used a driving simulation experiment to examine how different driving modes affect driving safety under different pre-drive states. The results suggested that, compared with manual driving, the hands-off mode had detrimental effects on drowsiness and attention behaviors in both driver groups, with adverse trends also observed for hazard perception

performance and driving performance. Given the adverse impact of the hands-off mode, non-nightshift workers took over more frequent than the nightshift workers in response to hazards. They maintained safer driving outcomes in the hands-off mode than their nightshift counterparts. In contrast, nightshift workers responded similarly across all driving modes but showed the highest collision probability in the hands-off mode.

Although L2 driving could reduce manual control demands, taking hands off the steering wheel may increase safety risks. These findings have important implications for regulating usage of L2 driving systems, especially for drivers with different pre-drive states. Alert drivers may be benefited from periodically placing their hands on the steering wheel after using hands-off mode for a certain period. However, for drivers who are already drowsy before driving, such as the nightshift workers, usage of L2 hands-off systems is not advisable, as they may be less capable to adopt safe response strategies. As the automobile industry is continuing to develop driver state monitoring and intervention systems in autonomous vehicles, the allowance of using L2 systems could be better adapted to drivers' states in the future.

Acknowledgement

We appreciate the support from National Natural Science Foundation of China with the Main Research Project on Machine Behavior and Human-Machine Collaborated Decision Making Methodology (Project Number: 72192820), and the Second Research Project on Human Behavior in Human-Machine Collaboration (Project Number: 72192824). We gratefully acknowledge the support of the China Scholarship Council. We would also like to express our sincere thanks to Prof. Natasha Merat at the University of Leeds and her team members, Rafael C. Gonçalves, Albert Solernou, and Yuwei Wang, for their invaluable support in completing this experiment.

References

- Aarts, L., & Van Schagen, I. (2006). Driving speed and the risk of road crashes: A review. *Accident Analysis & Prevention*, 38(2), 215–224. <https://doi.org/10.1016/j.aap.2005.07.004>
- Ahlström, C., Zemblys, R., Jansson, H., Forsberg, C., Karlsson, J., & Anund, A. (2021). Effects of partially automated driving on the development of driver sleepiness. *Accid Anal Prev*, 153, 106058. <https://doi.org/10.1016/j.aap.2021.106058>
- Åkerstedt, T., & Gillberg, M. (1990). Subjective and Objective Sleepiness in the Active Individual. *International Journal of Neuroscience*, 52(1–2), 29–37. <https://doi.org/10.3109/00207459008994241>
- American Psychological Association. (2018). *Drowsiness*. APA dictionary of psychology. <https://dictionary.apa.org/drowsiness>
- Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, 34(5), 502–508. <https://doi.org/10.1111/opo.12131>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Baldisserotto, F., Krejtz, K., Dominguez-Martinez, E., Duchowski, A. T., & Krejtz, I. (2026). Pupillary and attention dynamics in response to road hazard detection. *Transportation Research Part F: Traffic Psychology and Behaviour*, 118, 103543. <https://doi.org/10.1016/j.trf.2026.103543>
- Banks, S., & Dinges, D. F. (2007). Behavioral and Physiological Consequences of Sleep Restriction. *Journal of Clinical Sleep Medicine*, 03(05), 519–528. <https://doi.org/10.5664/jcsm.26918>
- Barde, M., & Barde, P. (2012). What to use to express the variability of data: Standard deviation or standard error of mean? *Perspectives in Clinical Research*, 3(3), 113. <https://doi.org/10.4103/2229-3485.100662>

- Ben-Bassat, T., & Shinar, D. (2011). Effect of shoulder width, guardrail and roadway geometry on driver perception and behavior. *Accident Analysis & Prevention*, 43(6), 2142–2152. <https://doi.org/10.1016/j.aap.2011.06.004>
- Berger, P. D., Maurer, R. E., & Celli, G. B. (2018). Experimental design (pp. 449-480). Cham: Springer International Publishing.
- Boguslavsky, D., Borowsky, A., Chassidim, H., & Hancock, P. A. (2025). Secondary task solutions to acute and chronic automation-induced underload. *Transportation Research Part F: Traffic Psychology and Behaviour*, 109, 1451–1469. <https://doi.org/10.1016/j.trf.2025.01.033>
- Borowsky, A., Oron-Gilad, T., Meir, A., & Parmet, Y. (2012). Drivers' perception of vulnerable road users: A hazard perception approach. *Accident Analysis & Prevention, Safety and Mobility of Vulnerable Road Users: Pedestrians, Bicyclists, and Motorcyclists*, 44(1), Article 1. <https://doi.org/10.1016/j.aap.2010.11.029>
- Cahour, B., Jean-Francois, F., & Arnaud, K. (2021). Driving a Partially Automated Car with the Hands On or Off the Steering Wheel: Users' Subjective Experiences. In N. L. Black, W. P. Neumann, & I. Noy (Eds.), *Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021)* (Vol. 221, pp. 583–592). Springer International Publishing. https://doi.org/10.1007/978-3-030-74608-7_71
- Cao, S., Samuel, S., Murzello, Y., Ding, W., Zhang, X., & Niu, J. (2022). Hazard Perception in Driving: A Systematic Literature Review. *Transportation Research Record: Journal of the Transportation Research Board*, 2676(12), 666–690. <https://doi.org/10.1177/03611981221096666>
- Carsten, O., Frank, L., Barnard, Y., Jamson, H., & Merat, N. (2012). Control task substitution in semiautomated driving: Does it matter what aspects are automated? - PubMed. *Human Factors*. <https://pubmed.ncbi.nlm.nih.gov/23156620/>

- Cantwell, S. J. (2021). The role of hazard perception in speed choice (Doctoral dissertation, The University of Waikato).
- Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1–8. <https://doi.org/10.1016/j.jebo.2011.08.009>
- Crundall, D., Chapman, P., Trawley, S., Collins, L., van Loon, E., Andrews, B., & Underwood, G. (2012). Some hazards are more attractive than others: Drivers of varying experience respond differently to different types of hazard. *Accident Analysis & Prevention*, 45, 600–609. <https://doi.org/10.1016/j.aap.2011.09.049>
- Damböck, D., Weissgerber, T., Kienle, M., & Bengler, K. (2013). Requirements for cooperative vehicle guidance. *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 1656–1661. <https://doi.org/10.1109/ITSC.2013.6728467>
- European Transport Safety Council. (2023). *UK permits hands-off Level 2 driver assistance system*. <https://etsc.eu/uk-permits-hands-off-level-2-driver-assistance-system/>
- Favarò, F. M., Seewald, P., Scholtes, M., & Eurich, S. (2019). Quality of control takeover following disengagements in semi-automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 64, 196–212. <https://doi.org/10.1016/j.trf.2019.05.004>
- Fisher, R. A. (1925). Statistical methods for research workers. Oliver and Boyd.
- Foong, R., Ang, K. K., Quek, C., Guan, C., & Phyoo Wai, A. A. (2015). An analysis on driver drowsiness based on reaction time and EEG band power. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 7982–7985. <https://doi.org/10.1109/EMBC.2015.7320244>
- Gershon, P., Mehler, B., & Reimer, B. (2023). Driver response and recovery following automation initiated disengagement in real-world hands-free driving. *Traffic*

Injury Prevention, 24(4), 356–361.
<https://doi.org/10.1080/15389588.2023.2189990>

Greenlee, E. T., DeLucia, P. R., & Newton, D. C. (2018). Driver Vigilance in Automated Vehicles: Hazard Detection Failures Are a Matter of Time. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 60(4), 465–476. <https://doi.org/10.1177/0018720818761711>

Greenlee, E. T., DeLucia, P. R., & Newton, D. C. (2024). Driver Vigilance Decrement is More Severe During Automated Driving than Manual Driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 66(2), 574–588. <https://doi.org/10.1177/00187208221103922>

Hays, W. L. (1973). *Statistics for the social sciences*, Vol. 410. Holt, Rinehart: Winston New York.

Horswill, M. S., & McKenna, F. P. (2004). Drivers' hazard perception ability: Situation awareness on the road. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 155–175). Aldershot, UK: Ashgate.

Huestegge, L., Skottke, E.-M., Anders, S., Müsseler, J., & Debus, G. (2010). The development of hazard perception: Dissociation of visual orientation and hazard processing. *Transportation Research Part F: Traffic Psychology and Behaviour*, 13(1), 1–8. <https://doi.org/10.1016/j.trf.2009.09.005>

Hussain, A., Large, D. R., & Harvey, C. (2025). Do Hands-On or Hands-Off Drivers Perform Better During a Level-2 Silent Failure? *Contemporary Ergonomics and Human Factors*.

Johns, M. W. (1991). A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale. *Sleep*, 14(6), 540–545. <https://doi.org/10.1093/sleep/14.6.540>

Kahana-Levy, N., Shavitzky-Golkin, S., Borowsky, A., & Vakil, E. (2019). The effects of repetitive presentation of specific hazards on eye movements in hazard

- perception training, of experienced and young-inexperienced drivers. *Accident Analysis & Prevention*, 122, 255-267.
- Kaplan, S., & Prato, C. G. (2012). Associating Crash Avoidance Maneuvers With Driver Attributes and Accident Characteristics: A Mixed Logit Model Approach. *Traffic Injury Prevention*, 13(3), 315–326. <https://doi.org/10.1080/15389588.2011.654015>
- Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. *Procedia Manufacturing*, 3, 2403–2409. <https://doi.org/10.1016/j.promfg.2015.07.499>
- Kosmidis, I., Kenne Pagui, E. C., Boiocchi, F., Sterzinger, P., Konis, K., & Sartori, N. (2025). *brglm2: Bias reduction in generalized linear models (R package version 0.x)*. <https://cran.r-project.org/web/packages/brglm2/index.html>
- Kosmidis, I., Kenne Pagui, E. C., & Sartori, N. (2020). Mean and median bias reduction in generalized linear models. *Statistics and Computing*, 30(1), 43–59. <https://doi.org/10.1007/s11222-019-09860-6>
- Large, D. R., Clark, L., Quandt, A., Burnett, G., & Skrypchuk, L. (2017). Steering the conversation: A linguistic exploration of natural language interactions with a digital assistant during simulated driving. *Applied Ergonomics*, 63, 53–61. <https://doi.org/10.1016/j.apergo.2017.04.003>
- Lee, D. K., In, J., & Lee, S. (2015). Standard deviation and standard error of the mean. *Korean Journal of Anesthesiology*, 68(3), 220. <https://doi.org/10.4097/kjae.2015.68.3.220>
- Lee, J., Gershon, P., Domeyer, J., Mehler, B., & Reimer, B. (2024). Understanding hands-off patterns while using Tesla Autopilot: Individual differences and their impact on hands-off duration. *THE 9TH INTERNATIONAL CONFERENCE ON DRIVER DISTRACTION AND INATTENTION*, 26–29.

- Lee, Y. T., Gonçalves, R. C., Öztürk, İ., Solernou, A., & Proctor, K. (2025). *Hazard perception in hands-off Level 2 driving: Results from a driving simulator study 4 during daytime and after dark lighting conditions*.
- Lehmann, E. L., & Romano, J. P. (2008). *Testing statistical hypotheses* (3. ed., correct. at 4. print). Springer.
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). *Package ‘emmeans.’* R package version, 1(3.2). <https://mirror.las.iastate.edu/CRAN/web/packages/emmeans/refman/emmeans.html>
- Ma, J., Gu, J., Jia, H., Yao, Z., & Chang, R. (2018). The relationship between drivers’ cognitive fatigue and speed variability during monotonous daytime driving. *Frontiers in psychology*, 9, 459.
- May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(3), 218–224. <https://doi.org/10.1016/j.trf.2008.11.005>
- Merat, N., & Jamson, A. H. (2013). The effect of three low-cost engineering treatments on driver fatigue: A driving simulator study. *Accident Analysis & Prevention*, 50, 8–15. <https://doi.org/10.1016/j.aap.2012.09.017>
- Merat, N., Seppelt, B., Louw, T., Engström, J., Lee, J. D., Johansson, E., Green, C. A., Katazaki, S., Monk, C., Itoh, M., McGehee, D., Sunda, T., Unoura, K., Victor, T., Schieben, A., & Keinath, A. (2019). The “Out-of-the-Loop” concept in automated driving: Proposed definition, measures and implications. *Cognition, Technology & Work*, 21(1), 87–98. <https://doi.org/10.1007/s10111-018-0525-8>
- Milech, D., Glencross, D., & Hartley, L. (1989). *Skills Acquisition by Young Drivers: Perceiving, Interpreting and Responding to the Driving Environment* (Issue MR4). <https://trid.trb.org/View/353420>

- Minderhoud, M. M., & Bovy, P. H. L. (2001). Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis & Prevention*, 33(1), 89–97. [https://doi.org/10.1016/S0001-4575\(00\)00019-1](https://doi.org/10.1016/S0001-4575(00)00019-1)
- Morad, Y., Lemberg, H., Yofe, N., & Dagan, Y. (2000). Pupillography as an objective indicator of fatigue. *Current Eye Research*, 21(1).
- National Sleep Foundation. (2023). *National Sleep Foundation's Drowsy Driving Survey*.
- Naujoks, F., Purucker, C., Neukum, A., Wolter, S., & Steiger, R. (2015). Controllability of Partially Automated Driving functions – Does it matter whether drivers are allowed to take their hands off the steering wheel? *Transportation Research Part F: Traffic Psychology and Behaviour*, 35, 185–198. <https://doi.org/10.1016/j.trf.2015.10.022>
- Noble, A. M., Miles, M., Perez, M. A., Guo, F., & Klauer, S. G. (2021). Evaluating driver eye glance behavior and secondary task engagement while using driving automation systems. *Accident Analysis & Prevention*, 151, 105959. <https://doi.org/10.1016/j.aap.2020.105959>
- Pan, H., Logan, D. B., Stephens, A. N., Payre, W., Wang, Y., Peng, Z., Qin, Y., & Koppel, S. (2025). Exploring the effect of driver drowsiness on takeover performance during automated driving: An updated literature review. *Accident Analysis & Prevention*, 216, 108023. <https://doi.org/10.1016/j.aap.2025.108023>
- Pan, H., Payre, W., Wang, Y., & Gao, Z. (2025). How do visual and cognitive non-driving related tasks affect drivers' visual attention and takeover performance in conditionally automated driving? *Journal of Transportation Safety & Security*, 17(1), 97–123. <https://doi.org/10.1080/19439962.2024.2368118>
- Radlmayr, J., Feldhütter, A., Frey, A., Jarosch, O., Naujoks, F., Weinbeer, V., & Bengler, K. (2019). *Drowsiness and fatigue in conditionally automated driving – Towards an integrative framework*. Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2018 Annual Conference.

- Robinson, C. D., Tomek, S., & Schumacker, R. E. (2013). *Tests of Moderation Effects: Difference in Simple Slopes versus the Interaction Term*. 39(1).
- SAE International. (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles—SAE International*.
https://www.sae.org/standards/content/j3016_202104/
- Samuel, S., Yamani, Y., & Fisher, D. (2020). Understanding drivers' latent hazard anticipation in partially automated vehicle systems. *International Journal of Human Factors and Ergonomics*, 7, 282.
<https://doi.org/10.1504/IJHFE.2020.110093>
- Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110, 104038. <https://doi.org/10.1016/j.jml.2019.104038>
- Smit, A. S., Eling, P. A. T. M., Hopman, M. T., & Coenen, A. M. L. (2005). Mental and physical effort affect vigilance differently. *International Journal of Psychophysiology*, 57(3), 211–217.
<https://doi.org/10.1016/j.ijpsycho.2005.02.001>
- Smith, S. S., Horswill, M. S., Chambers, B., & Wetton, M. (2009). Hazard perception in novice and experienced drivers: The effects of sleepiness. *Accident Analysis & Prevention*, 41(4), 729–733. <https://doi.org/10.1016/j.aap.2009.03.016>
- Soulmana, B., Boukebbab, S., & Boulahlib, M. S. (2021). Hand position on steering wheel during fatigue and sleepiness case: Driving simulator. *Advances in transportation studies*, 53, 69-84.
- UK Government. (n.d.). *Speed limits*. Retrieved <https://www.gov.uk/speed-limits>
- Underwood, G., Ngai, A., & Underwood, J. (2013). Driving experience and situation awareness in hazard detection. *Safety Science*, 56, 29–35.
<https://doi.org/10.1016/j.ssci.2012.05.025>

- Van Der Horst, R., & De Ridder, S. (2007). Influence of roadside infrastructure on driving behavior: Driving simulator study. *Transportation Research Record*, 36–44.
- Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer New York. <https://doi.org/10.1007/978-0-387-21706-2>
- Vogelpohl, T., Kühn, M., Hummel, T., & Vollrath, M. (2019). Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. *Accident Analysis & Prevention*, 126, 70–84. <https://doi.org/10.1016/j.aap.2018.03.013>
- Wan, Y., & Sarter, N. (2022). Attention Limitations in the Detection and Identification of Alarms in Close Temporal Proximity. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 001872082110639. <https://doi.org/10.1177/00187208211063991>
- Watling, C. N., & Home, M. (2022). Hazard perception performance and visual scanning behaviours: The effect of sleepiness. *Transportation Research Part F: Traffic Psychology and Behaviour*, 90, 243–251. <https://doi.org/10.1016/j.trf.2022.08.020>
- Weaver, B. W., DeLucia, P. R., & Jupe, J. (2023). Factors That Affect Drivers' Perception of Closing and an Immediate Hazard. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 65(1), 166–181. <https://doi.org/10.1177/00187208211009028>
- Winn, B., & Whitaker, D. (1994). *Factors Affecting Light-Adapted Pupil Size in Normal Human Subjects*. 35(3).
- Yang, J., Liu, C., Chu, P., Wen, X., & Zhang, Y. (2021). Exploration of the Relationships between Hazard Perception and Eye Movement for Young Drivers. *Journal of Advanced Transportation*, 2021, 1–10. <https://doi.org/10.1155/2021/6642999>
- Zangi, N., Srour-Zreik, R., Ridel, D., Chassidim, H., & Borowsky, A. (2022). Driver distraction and its effects on partially automated driving performance: A driving

- simulator study among young-experienced drivers. *Accident Analysis & Prevention*, 166, 106565. <https://doi.org/10.1016/j.aap.2022.106565>
- Zhang, N., Fard, M., Davy, J. L., Parida, S., & Robinson, S. R. (2023). Is driving experience all that matters? Drivers' takeover performance in conditionally automated driving. *Journal of Safety Research*, 87, 323–331. <https://doi.org/10.1016/j.jsr.2023.08.003>
- Zhang, W., Carsten, O., Goodridge, C. M., & Zhang, W. (2025). Effects of adopting voice assistants countermeasures on alerting drowsy drivers across age. *Applied Ergonomics*, 129, 104616. <https://doi.org/10.1016/j.apergo.2025.104616>

Appendix

Table A.1. Experimental drives and hazard arrangements.

Driving Test	Hazard Sequence	Hazard Type	Hazard Attribute	Hazard Direction	Appear Distance (m)
Driving test 1	Hazard 1	Pedestrian	Hidden unmaterialized	Left	2,900
	Hazard 5	Car	Hidden unmaterialized	Right	7,000
	Hazard 4	Car	Hidden unmaterialized	Left	8,085
	Hazard 2	Pedestrian	Hidden unmaterialized	Right	11,700
	Hazard 6	Car	Hidden unmaterialized	Right	14200
	Hazard 3	Pedestrian	Hidden materialized	Left	16,000
Driving test 2	Hazard 3	Pedestrian	Hidden unmaterialized	Left	2,900
	Hazard 1	Pedestrian	Hidden unmaterialized	Right	7,000
	Hazard 5	Car	Hidden unmaterialized	Right	8,085
	Hazard 6	Car	Hidden unmaterialized	Left	11180
	Hazard 2	Pedestrian	Hidden unmaterialized	Right	14,200
	Hazard 4	Car	Hidden materialized	Left	16,000
Driving test 3	Hazard 6	Car	Hidden unmaterialized	Right	3,169
	Hazard 2	Pedestrian	Hidden unmaterialized	Left	4,800
	Hazard 3	Pedestrian	Hidden unmaterialized	Right	7,880
	Hazard 4	Car	Hidden unmaterialized	Right	10,105
	Hazard 5	Car	Hidden unmaterialized	Left	12,600
	Hazard 1	Pedestrian	Hidden materialized	Left	14,400

Table A.2. Omnibus tests (Step 1): effects of driver group and driving mode on continuous dependent variables

Drowsiness indicators	After-drive KSS scores					PD				
	df ₁	df ₂	F	<i>p</i>	η_p^2	df ₁	df ₂	F	<i>p</i>	η_p^2
Driver group	1	36	6.091	0.018	0.145	1	36.3	1.101	0.301	0.02
Driving mode	2	72	5.351	0.007	0.129	2	477.52	3.941	0.020	0.02
Driver group $\times$ Driving mode	2	72	0.242	0.786	0.007	2	477.52	1.705	0.183	0.007
Hazard perception performance	TTFF									
	df ₁	df ₂	F	<i>p</i>	η_p^2					
Driver group	1	35.82	25.501	< .001	0.420					
Driving mode	2	480.51	0.669	0.513	0.003					
Driver group $\times$ Driving mode	2	480.51	0.411	0.664	0.002					
Gaze behaviors	P _{Infotainment Display}					P _{Driving Road}				
	df ₁	df ₂	F	<i>p</i>	η_p^2	df ₁	df ₂	F	<i>p</i>	η_p^2
Driver group	1	36	0.095	0.760	0.003	1	36	1.215	0.270	0.000
Driving mode	2	72	7.656	<0.001	0.180	2	72	12.288	<0.001	0.210
Driver group $\times$ Driving mode	2	72	0.548	0.580	0.010	2	72	2.262	0.323	0.030
	P _{Instrument Panel}									
	df ₁	df ₂	F	<i>p</i>	η_p^2					
Driver group	1	36	0.039	0.843	0.000					
Driving mode	2	72	10.783	0.005	0.210					
Driver group $\times$ Driving mode	2	72	0.128	0.938	0.002					
	SD of horizontal gaze					SD of vertical gaze				
	df ₁	df ₂	F	<i>p</i>	η_p^2	df ₁	df ₂	F	<i>p</i>	η_p^2
Driver group	1	36	0.011	0.918	0.003	1	36	2.365	0.133	0.060
Driving mode	2	72	22.114	< .001	0.380	2	72	43.936	< .001	0.550
Driver group $\times$ Driving mode	2	72	0.232	0.794	0.006	2	72	0.084	0.919	0.002
Driving performance	TTC _{response}									
	df ₁	df ₂	F	<i>p</i>	η_p^2					
Driver group	1	33.01	0.423	0.520	0.010					
Driving mode	2	59.77	2.587	0.084	0.080					
Driver group $\times$ Driving mode	2	59.77	2.272	0.112	0.070					

Note. η_p^2 indicates the partial eta square, determining the effect size. Degrees of freedom for fixed effects were approximated using Satterthwaite's method, which can yield non-integer denominator degrees of freedom.

Table A.3. Omnibus tests (Step 1): effects of driver group and driving mode on binary dependent variables

Predictors	Coefficient (β)	SE	Z	p	OR	CI-L	CI-U
<i>Hazard identification</i>							
Intercept	3.241	0.548	5.917	< .001	25.571	8.738	74.830
Driver Group [Nightshift]	-1.230	0.640	-1.922	0.033	0.292	0.083	1.025
Driving Mode [L2 hands-on]	0.348	0.846	0.411	0.681	1.416	0.270	7.431
Driving Mode [Manual]	1.946	1.528	1.274	0.203	7.000	0.351	139.800
Driving Group [Nightshift] *	0.229	0.996	0.230	0.818	1.257	0.179	8.857
Driving Mode [L2 hands-on]							
Driving Group [Nightshift] *	-1.661	1.604	-1.035	0.301	0.190	0.008	4.409
Driving Mode [Manual]							
<i>Response</i>							
Intercept	-0.339	0.342	-0.991	0.322	0.712	0.364	1.393
Driver Group [Nightshift]	-0.306	0.489	-0.626	0.531	0.736	0.282	1.920
Driving Mode [L2 hands-on]	-0.389	0.331	-1.174	0.240	0.678	0.354	1.297
Driving Mode [Manual]	-0.691	0.340	-2.033	0.042	0.501	0.257	0.975
Driving Group [Nightshift] *	0.028	0.477	0.059	0.953	1.028	0.404	2.620
Driving Mode [L2 hands-on]							
Driving Group [Nightshift] *	0.691	0.478	1.446	0.148	1.995	0.782	5.089
Driving Mode [Manual]							
<i>Collision</i>							
Intercept	-1.946	0.694	-2.805	0.005	0.143	0.037	0.556
Driver Group [Nightshift]	1.435	0.840	1.708	0.088	4.200	0.809	21.795
Driving Mode [L2 hands-on]	0.395	0.920	0.430	0.667	1.485	0.245	9.005
Driving Mode [Manual]	-1.718	1.625	-1.057	0.290	0.179	0.007	4.337
Driving Group [Nightshift] *	-0.615	1.145	-0.538	0.591	0.540	0.057	5.094
Driving Mode [L2 hands-on]							
Driving Group [Nightshift] *	-0.284	1.904	-0.149	0.881	0.753	0.018	31.413
Driving Mode [Manual]							

Note. Probability distribution: Binomial (Logit). Hazard identified = 1; Not identified = 0; Responded = 1; Did not respond = 0; Collided = 1; Did not collide = 0.