Advances in Human Factors of Transportation, Vol. 224, 2026, 351–360
https://doi.org/10.54941/ahfe1007863

Physiological Assessment of Driver Trust in Automated Vehicles Under Distinct Driving Styles

Yizheng Li[1], Zhilin Hu[1], Andrew P. Owens[2], Karl Proctor[2], Lisa Dorn[1] and Yifan Zhao[1]

[1]Faculty of Engineering and Applied Sciences, Cranfield University, Bedford, UK
[2]Jaguar Land Rover Research, National Automotive Innovation Centre, Coventry, UK

ABSTRACT

With the rapid advancement of automated vehicle (AV) technology, drivers' willingness to rely on AV systems has become a decisive factor in their broader deployment. Among the many psychological determinants, drivers' trust has been consistently identified as a key condition influencing the acceptance and appropriate use of AVs. Consequently, understanding how trust evolves during real-time interaction with an AV is increasingly critical for determining whether drivers can engage with the system safely and effectively. Current research on trust assessment often relies on subjective measures, yielding limited temporal resolution and vulnerability to subjective bias. This study proposes a multimodal, event-driven experimental framework to investigate how different AV driving styles affect driver trust through objective physiological signals. First, a driving simulation experiment was designed with three distinct driving styles: a) a baseline scenario featuring smooth, normative driving to establish physiological reference physiological states; b) a hesitant scenario characterised by cautious and delayed decision-making; and c) an aggression condition characterised by late braking and assertive, high-speed cornering. Each participant experienced all conditions in counterbalanced order. Second, participants' physiological responses were recorded using a synchronised multi-sensor setup, including electroencephalography (EEG) and eye-tracking. Subjective trust ratings were collected after each scenario to serve as the ground truth for trust evaluation. Finally, the collected signals were integrated to analyse the correlation between trust level and physiological features. The results suggest that aggressive and hesitant driving styles elicit distinguishable subjective, neural, and attentional responses, indicating different pathways of distrust. These findings provide preliminary evidence supporting the feasibility of physiology-based approachesfor assessing driver trust in automated driving.

Keywords: Automated vehicle, Driver's trust, Driving styles, Physiological signals, Driving simulation

INTRODUCTION

Trust plays a critical role in human interaction with automated driving systems. Appropriate levels of trust are essential for effective supervision, timely intervention, and long-term acceptance of Level 3 (L3) automated vehicles (AVs), where the driver must remain available to resume control upon request (Jian et al., 2000). Over-trust leads to misuse or abuse, while

Received March 20, 2026; Revised April 26, 2026; Accepted May 15, 2026; Available online July 20, 2026

under-trust results in the disuse of beneficial technologies. Importantly, trust in AV is not a static state (Raats et al., 2020); rather, it evolves dynamically in response to system and situational context, particularly during atypical or unexpected driving events.

Existing research on trust in AV has predominantly relied on subjective self-report measures and drivers' behavioural features, such as trust scales (Zhang et al., 2019) and takeover frequency (Yi et al., 2023). While valuable for assessing overall attitudes, subjective methods suffer from subjective bias and often fail to capture the real-time trust level, and the behaviours often confound trust with task decisions or driving workload. As a result, the underlying cognitive mechanisms of trust remain insufficiently understood. To address this limitation, physiological sensing modalities, including galvanic skin response, heart rate, eye tracking and electroencephalography (EEG), become valuable methods for capturing continuous indicators of trust. Prior studies have shown that gaze behaviour (Cai et al., 2025) can reflect visual attention and situational awareness, while EEG features are sensitive to cognitive workload and risk processing (Ma et al., n.d.). However, integrating these modalities into AV trust research poses significant challenges due to their low signal-to-noise ratios and motion artefacts.

Moreover, most existing studies on trust design AV system failures with binary scenarios (trustworthy vs. untrustworthy) (Tan and Hao, 2023). However, real-world AV trust might be caused by different dimensions of vehicle-related factors. For instance, AV aggressive behaviours may induce fear and perceived danger, whereas hesitant behaviours may signal incompetence without immediate physical threat (Bellem et al., 2018; Dillen et al., 2020). This study hypothesises that these distinct behaviours impact trust through different psychological pathways, yet their neural correlates remain underexplored.

In this work, we propose a multimodal, event-driven experimental framework to investigate trust in AV, integrating EEG, eye-tracking, and subjective measures during distinct driving styles. Using a moderate fidelity driving simulator, we designed three automated driving scenarios—Baseline, Aggressive, and Hesitant—implemented on the same driving route. This research aims to identify whether distinct eye movements and neural signatures can be associated with trust in distinct driving styles. This work lays the foundation for future large-sample studies aimed at uncovering the relationships between automated driving behaviour, physiological responses, and human trust. Ultimately, such insights may contribute to the development of trust-aware automated driving systems that better align with human expectations and cognitive states.

METHODOLOGY

Participant

A total of 13 participants (Mean Age = 25.4 ± 3.2 years) were recruited for this study. All participants possessed a valid driving licence with at least 2 years of driving experience. The study protocol was approved by the Cranfield University Research Ethics System (Reference: CURE/24703/2025).

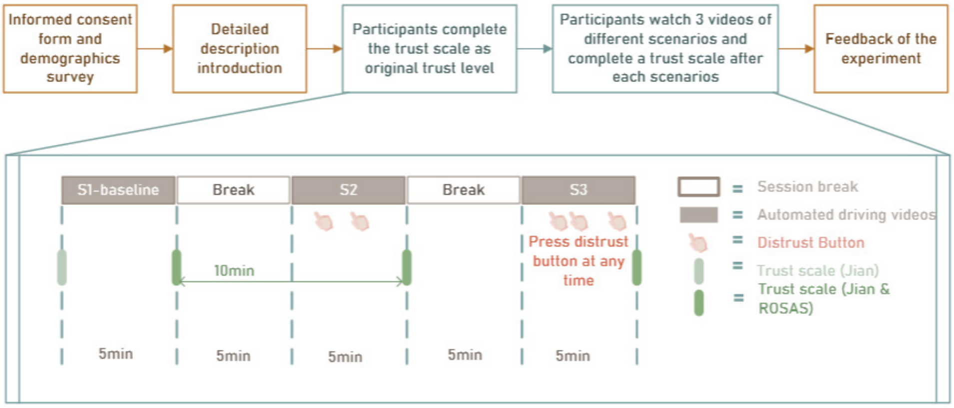


Figure 1: Experimental procedure.

Experimental Apparatus and Procedure

The experiment was conducted in a laboratory setting. The driving simulator consists of a three-screen panoramic display (27-inch each, 7680*1440 resolution), a Thrustmaster RS300 (force-feedback steering wheel and pedal system), and a GPU with 32GB of RAM. The driving simulator was constructed using the BeamNG.tech software platform, which logged vehicle state variables, including speed, distance and destruction, and participants' input at 120Hz. The eye tracker captured head position, head rotation, gaze origin and direction at 120Hz. EEG signals were recorded using an EEGO system with 24 electrodes arranged according to the international 10–20 system.

The experiment followed a structured and route-controlled procedure, illustrated in Figure 1. Upon arrival, participants provided informed consent and completed a demographic questionnaire, followed by an initial trust assessment to establish a baseline trust level based on Jian's trust scale (Jian et al., 2000). After a standardised task briefing, participants experienced three automated driving scenarios, with short rest breaks between scenarios to mitigate fatigue effects. During each scenario, participants were instructed to continuously monitor the automated driving and were allowed to press a designated distrust button at any time they felt discomfort or loss of confidence. After each scenario, participants completed Jian's trust scale together with a ROSAS scale (Carpinella et al., 2017). A final debriefing and feedback session concluded the experiment.

The three scenarios (Baseline, Aggressive, Hesitant) were implemented to impose varying AV driving styles on the same driving route (Table 1). The Baseline scenario exhibited moderate and stable longitudinal and lateral control, serving as a neutral reference condition without salient anomalous events. The Aggressive scenario introduced assertive and risk-prone automated behaviours, including aggressive cornering and brief acceleration bursts, designed to challenge perceived safety and controllability. The Hesitant scenario emphasised conservative and temporally inefficient control, such as prolonged stopping and early braking, intended to induce uncertainty without explicit safety violations. Participants experienced the Baseline Scenario first, followed by the other two in random order to minimise the probable effect on the trust dynamic.

Table 1: Experimental scenario design.

Scenarios	Speed & Aggression Level	Events
Baseline	Moderate	None
Aggressive	High	• Curb contact • Cornering overshoot • Unchecked acceleration
Hesitant	Low	• Prolonged stopping • Slow creeping • Premature braking

Data Collection and Processing

Multimodal signals were temporally aligned using a unified timestamping framework to enable synchronised analysis across drivers' physiological modalities and events during automated driving. Discrete event markers, including scenario-specific events, destruction state and distrust-button-pressed moments, were logged with system timestamps and served as anchors for event-related analyses.

EEG data were preprocessed following standard procedures to ensure signal quality and comparability across conditions. Signals were band-pass filtered between 1–40 Hz and visually inspected for bad channels and noisy segments. Independent Component Analysis (ICA) was used to identify and remove components associated with eye blinking and muscle movements.

Cleaned EEG signals were segmented into time windows centred on the marked events from –4000 ms to 0 ms. Time–frequency analysis was conducted using short-time Fourier transform (STFT) with a sliding window of 1 s and a step size of 0.5 s, signals averaged during each window to enable temporally resolved spectral estimation within each epoch.

Non-neural data streams, eye-tracking, were temporally aligned and epoched to correspond with the cleaned EEG segments. All preprocessing, segmentation, and feature extraction procedures were applied consistently across scenarios and participants.

RESULT AND DISCUSSION

Subjective Trust Analysis

The comparative analysis of subjective trust measures revealed differential effects across driving styles (Figure 2). An initial trust score (Origin) was collected before the experiment to characterise participants' initial trust toward automated driving, whereas Baseline (B) reflects trust after experiencing normal automated driving within the simulator. The Aggressive scenario significantly reduced overall trust compared to the Baseline

($p < 0.01$), whereas the Hesitant scenario did not produce a statistically significant decrease in trust ($p = 0.77$). Consistently, relative trust deviation (ΔTrust) indicated a larger trust reduction under aggressive driving than under hesitant driving ($p < 0.05$) (Figure 2b). Analysis of the ROSAS-Competence subscale showed (Figure 2c, 2d) that perceived system competence declined more sharply in the Aggressive condition than in the Hesitant condition ($\Delta M_A = -1.42$ vs. $\Delta M_H = -0.50$, $p < 0.05$). In addition, subjective discomfort increased substantially in the aggressive scenario relative to hesitation ($\Delta M_A = +1.85$ vs. $\Delta M_H = +0.27$, $p < 0.01$).

Taken together, these findings suggest that assertive, safety-critical driving behaviours are associated with stronger reductions in perceived competence and greater affective discomfort than temporally inefficient but conservative behaviours, indicating differentiated impacts on driver trust.

Table 2: Eye-tracking metrics across experimental conditions.

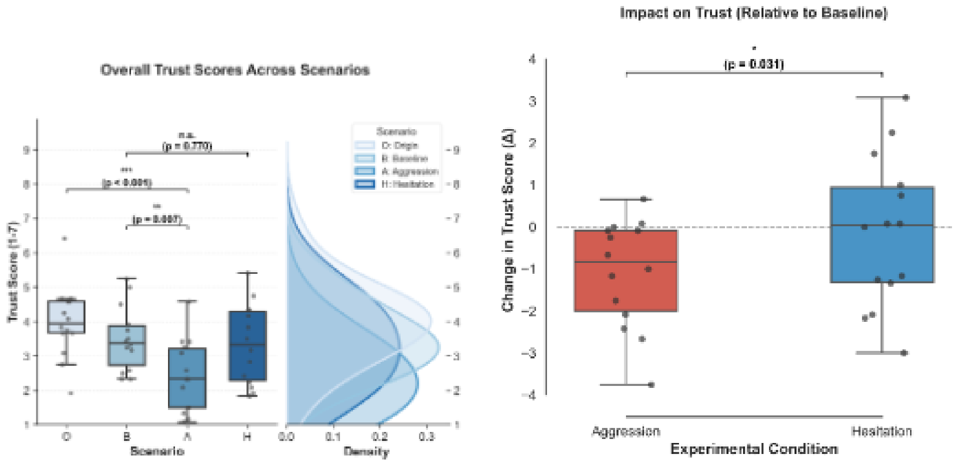
Condition	Avg Fixation Duration (ms)	Long Fixation Count (>400ms)	Total Long Fixation Time (s)	Avg Saccade Amplitude (Norm)	Large Saccade Prob (>0.3 Width)
Baseline	600.42	114.9	133.82	0.130	9.68%
Aggressive	658.47	134.4	166.03	0.121	**7.19%▲**
Hesitant	549.33	**200.6 ▲**	**229.32 ▲**	0.130	9.21%

Eye-Tracking Analysis

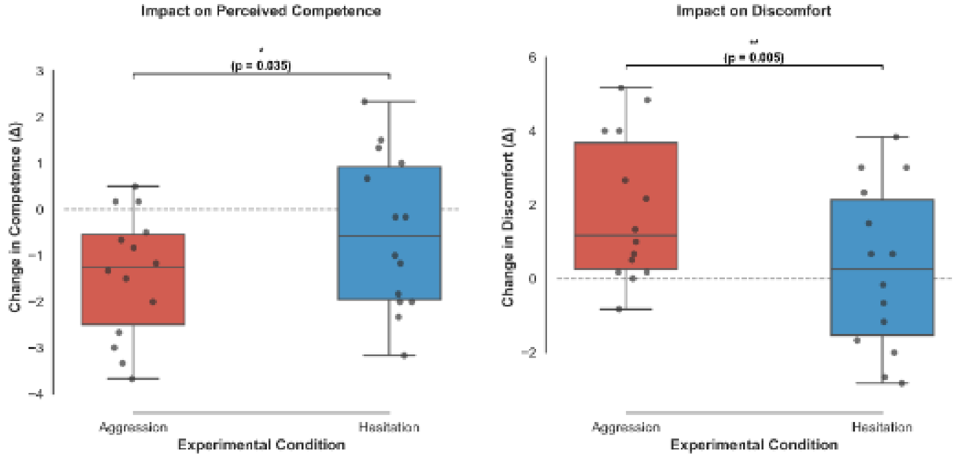
Global eye-tracking metrics were computed for each driving scenario to characterise overall visual attention patterns. As shown in Table 2, differences were observed across conditions in fixation-related and saccade- related measures. Relative to the Baseline condition, the Aggressive scenario was associated with a reduction in average saccade amplitude (0.130 to 0.121) and a lower probability of large saccades (9.68% to 7.19%). This pattern suggests a contraction of gaze dispersion and reduced exploratory eye movements under aggressive automated behaviour (Louw and Merat, 2017). In contrast, the Hesitant condition showed a marked increase in prolonged fixation behaviour. The number of long fixations (>400 ms) increased substantially (114.9 to 200.6), accompanied by a corresponding rise in total long fixation time (133.82 s to 229.32 s). Average saccade amplitude remained comparable to Baseline, and the probability of large saccades showed minimal change.

The observed eye-tracking patterns suggest that variations in automated driving style are accompanied by distinct visual attention strategies that may relate to changes in driver trust. Aggressive driving behaviour appears to induce sustained fixations with reduced large saccades, potentially reflecting heightened supervisory monitoring when confidence in the automation is challenged. Hesitant driving, by contrast, was characterised by prolonged fixation behaviour and extended total fixation time, which may reflect reduced trust in the system's decision-making consistency (Wang et al., 2024).

Taken together, these results indicate that different forms of trust decrease may manifest through differentiated gaze dynamics.



(a) Absolute Trust Scores (b) Trust Deviation (ΔTrust)

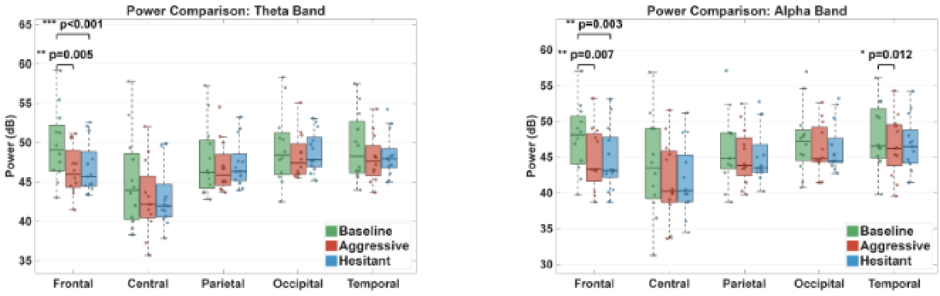


(c) Competence Deviation (ΔCompetence) (d) Discomfort Deviation (ΔDiscomfort)

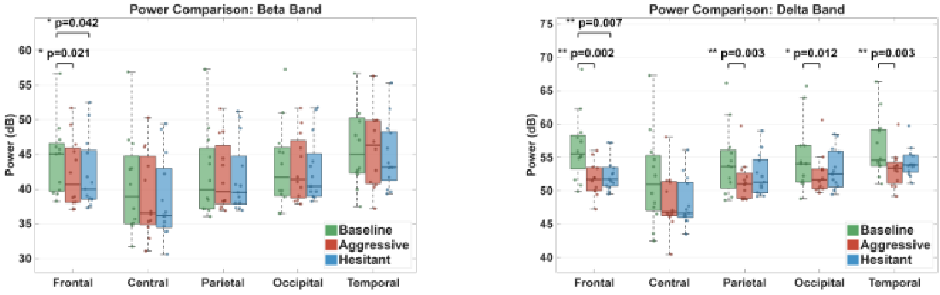
Figure 2: Subjective trust and human–machine perception across driving scenarios. (a) Absolute Jian Trust scores across Origin (pre-experiment), Baseline (normal automated driving), Aggressive, and Hesitant conditions. (b–d) Baseline-relative deviations in overall trust, perceived competence, and subjective discomfort. Red and blue denote Aggressive and Hesitant conditions, respectively. Asterisks indicate paired t-test significance (* $p < 0.05$, ** $p < 0.01$).

EEG Time-Frequency Analysis

EEG data were analysed using event-aligned segments to examine neural patterns (Figure 3). Relative to the Baseline condition, both Aggressive and Hesitant scenarios were associated with observable power reductions in the activity of multiple frequency bands. The lower-frequency bands (Delta and Theta) exhibit the most robust and widespread attenuation, while the higher-frequency bands (Alpha and Beta) demonstrate a milder suppression. Statistical comparisons indicated that these effects were consistently observed in the frontal region, where power decreased during both conditions in the four bands (all $p < 0.05$). In addition, the Aggression scenario exhibited broader spectral modulation, with significant power reductions in parietal, occipital, and temporal regions within the Delta band ($p \leq 0.012$), as well as temporal suppression in the Alpha band ($p = 0.012$).



(a) Theta band (4–8 Hz) (b) Alpha band (8–13 Hz)



(c) Beta band (13–30 Hz) (d) Delta band (0.5–4 Hz)

Figure 3: Normalised EEG power across frequency bands and brain regions.

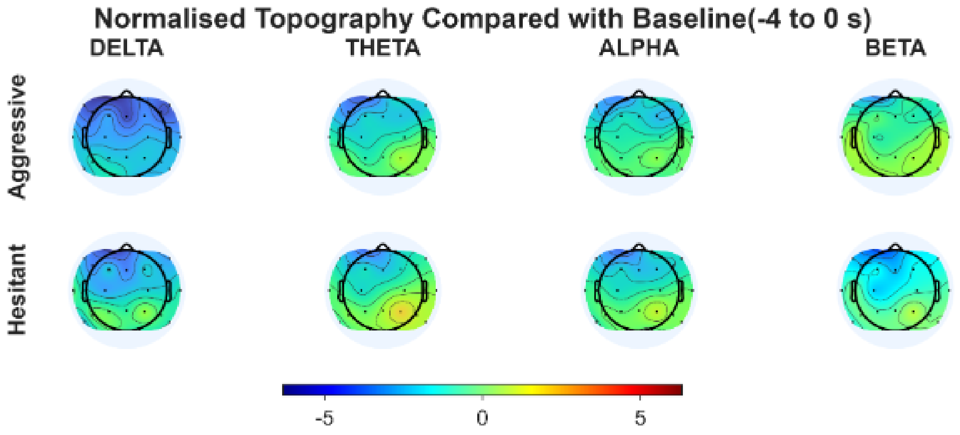


Figure 4: Normalised topography.

Figure 4 presents a more intuitive baseline-normalised topography across four frequency bands during [−4, 0]s. Similar to Figure 3, both the Aggression and Hesitation scenarios are associated with negative power deviations primarily distributed over the frontal region, particularly in the Delta band.

Moreover, the frontal power suppression appeared more pronounced over the left hemisphere compared to the right. It indicates that the frontal region has been widely associated with increased cognitive workload and the allocation of executive control during demanding tasks (Friedman and Robbins, 2022). A comparison between the two scenarios suggests subtle topographical differences. The Hesitation condition appears to exhibit relatively lower power at the frontal region, whereas the Aggression condition shows a more distributed pattern without marked frontal concentration.

These observations suggest that distrust induced by different automated driving styles may be associated with distinguishable spatial–spectral EEG dynamics. The topographical patterns further demonstrate the potential of characterising neural responses to variations in drivers' trust. Nevertheless, given the exploratory design and limited sample size, the present findings should be considered preliminary and primarily indicative of the methodological applicability of the proposed event-related EEG analysis.

Multimodal Convergence Analysis

Across subjective, ocular, and neural measures, a convergent pattern emerged linking automated driving style to differentiated trust responses. Subjective reports indicated that aggressive driving elicited greater reductions in perceived competence and increased discomfort compared to hesitant behaviour. In parallel, eye-tracking results showed sustained fixation patterns under aggressive driving, suggesting heightened scene monitoring, while EEG analyses revealed power suppression at the frontal region across multiple frequency bands. Their directional changes were theoretically aligned, collectively indicating elevated cognitive vigilance and reduced trust under an aggressive driving style.

Importantly, the present study does not claim direct statistical coupling between EEG bands and specific gaze metrics. Rather, the multimodal evidence demonstrates cross-modal consistency: behavioural attention allocation and neural spectral modulation shifted in patterns that corresponded with subjective trust decrease. Future work aims at joint feature analysis, which will be necessary to quantitatively establish functional relationships between neural activity, visual attention, and perceived trust.

CONCLUSION

This paper presents a multimodal experimental framework for trust in AV using synchronised behavioural and physiological signals. By integrating event-related scenario design with eye-tracking, EEG, and subjective measures in a high-fidelity driving simulator, the study demonstrates the feasibility of capturing trust-related physiological responses at both global and event-related levels.

Although the current results are preliminary due to the limited sample size, descriptive trends in eye-tracking and EEG time–frequency analyses suggest that different automated driving styles lead to distinct patterns of visual attention and neural activity. Importantly, the proposed methodology provides a scalable foundation for future large-sample studies.

Future work will extend this framework with increased participant numbers, event-level gaze analysis, and refined EEG features to further examine the relationship between drivers' trust and the driving styles of AV.

ACKNOWLEDGMENT

This study was funded by the EPSRC Industrial CASE (ICASE) studentship and Jaguar Land Rover (voucher reference: 240102).

REFERENCES

Bellem, H., Thiel, B., Schrauf, M. and Krems, J. F. (2018), 'Comfort in automated driving: An analysis of preferences for different automated driving styles and their dependence on personality traits', *Transportation research part F: traffic psychology and behaviour* **55**, 90–100.

Cai, Y., Demir, M., Sasangohar, F. and Zare, M. (2025), 'Investigating shifts in driver attention and trust across manual and autonomous driving: Insights from eye-tracking metrics', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **69**(1), 833–840.

Carpinella, C. M., Wyman, A. B., Perez, M. A. and Stroessner, S. J. (2017), The robotic social attributes scale (rosas): Development and validation, *in* 'Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot

Interaction', HRI '17, Association for Computing Machinery, New York, NY, USA, pp. 254–262.

Dillen, N., Ilievski, M., Law, E., Nacke, L. E., Czarnecki, K. and Schneider, O. (2020), Keep calm and ride along: Passenger comfort and anxiety as physiological responses to autonomous driving styles, *in* 'Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems', CHI '20, Association for Computing Machinery, New York, NY, USA, pp. 1–13.

Friedman, N. P. and Robbins, T. W. (2022), 'The role of prefrontal cortex in cognitive control and executive function', *Neuropsychopharmacology* **47**(1), 72– 89.

Jian, J.-Y., Bisantz, A. M. and Drury, C. G. (2000), 'Foundations for an empirically determined scale of trust in automated systems', *International Journal of Cognitive Ergonomics* **4**(1), 53–71.

Louw, T. and Merat, N. (2017), 'Are you in the loop? using gaze dispersion to understand driver visual attention during vehicle automation', *Transportation Research Part C: Emerging Technologies* **76**, 35–50.

Ma, S., Yan, X., Billington, J., Merat, N. and Markkula, G. (n.d.), 'Cognitive load during driving: EEG microstate metrics are sensitive to task difficulty and predict safety outcomes', **207**, 107769.

Raats, K., Fors, V. and Pink, S. (2020), 'Trusting autonomous vehicles: An interdisciplinary approach', *Transportation Research Interdisciplinary Perspectives* 7, 100201.

Tan, H. and Hao, Y. (2023), 'How does people's trust in automated vehicles change after automation errors occur? an empirical study on dynamic trust in automated driving', *Human Factors and Ergonomics in Manufacturing & Service Industries* **33**(6), 449–463.

Wang, K., Hou, W., Ma, H. and Hong, L. (2024), 'Eye-tracking characteristics: Unveiling trust calibration states in automated supervisory control tasks', *Sensors* **24**(24), 7946.

Yi, B., Cao, H., Song, X., Wang, J., Guo, W. and Huang, Z. (2023), 'Measurement and real-time recognition of driver trust in conditionally automated vehicles: Using multimodal feature fusions network', *Transportation research record* **2677**(8), 311–330.

Zhang, T., Tao, D., Qu, X., Zhang, X., Lin, R. and Zhang, W. (2019), 'The roles of initial trust and perceived risk in public's acceptance of automated vehicles', *Transportation research part C: emerging technologies* **98**, 207–220.