Exploring How Physio-Psychological States Affect Drivers' Takeover Performance in Conditional Automated Vehicles

Ange Wang[a,b], Jiyao Wang[c], Chunxi Huang[c], Dengbo He[a,b,c,d*], Hai Yang[a,b]

[a] *Thrust of Intelligent Transportation, Systems Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China*

[b] *Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China*

[c] *Thrust of Robotics and Autonomous Systems, Systems Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China*

[d] *HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen*

Email Addresses:

Ange Wang: awang324@connect.hkust-gz.edu.cn

Jiyao Wang: jwang297@connect.hkust-gz.edu.cn

Chunxi Huang: tracyhuang@hnu.edu.cn

Dengbo He: dengbohe@hkust-gz.edu.cn

Hai Yang: cehyang@ust.hk

Abstract

Although driving automation is promised to improve driving safety, drivers are still required to take over the control of the vehicles in case of emergency. Estimating drivers' takeover performance serves as the basis for adaptive driving automation and takeover request (TOR) to ensure driving safety. However, although algorithms have been proposed to estimate drivers' takeover performance through physiological and eye-tracking measures, the complex interrelationships between these metrics and driver behavior, as well as the interactions among the metrics themselves, are not fully understood. To answer this question, a driving simulation experiment involving 42 participants was conducted. Drivers experienced three types of takeover scenarios requested by TOR while driving a conditionally automated vehicle. Drivers' physiological, eye-tracking metrics and psychological states, as imposed by several non-driving-related tasks were collected. A structural equation model was used to explore the interactions among physiological metrics (i.e., cardiac activity, respiratory activity, electrodermal activity), eye-tracking metrics, psychological states (i.e., trust in driving automation and perceived workload), and variations in takeover time and takeover quality. The results showed that trust was positively associated with takeover quality, while workload was positively associated with takeover time. Additionally, physiological and eye-tracking metrics were indirectly associated with takeover quality via psychological states. This study reveals the hierarchical relationship among takeover performance-related variables and provides insights for designing driver monitoring systems aimed at estimating takeover performance in vehicles with driving automation and adaptive driving automation to improve driving safety.

Keywords

TOR, Physiological signals, Psychological states, Structural Equation Model, Driving Automation.

1 Introduction

Although autonomous driving technology is promised to enhance traffic safety and driving comfort, current automated driving systems still require driver interventions (SAE Levels 2 or 3). This brings critical challenges to driving safety, especially when drivers are asked to take over vehicle control by the driving automation systems in emergent situations, a scenario labeled as "system-initiated" control transitions. The system-initiated takeovers can occur when the environmental conditions are beyond the system's capabilities.

Given the importance of takeover performance for driving safety, previous studies have explored machine-learning models to predict driver takeover performance, intending to provide adaptive strategies to support drivers in takeover scenarios. For example, in a simulator study, Du et al. (2020) found significant changes in drivers' heart rate (HR), galvanic skin response (GSR), and eye-tracking metrics across different takeover scenarios and they successfully predicted drivers' takeover performance using these indicators (Du et al., 2020). Besides, Zhu et al. (2023) proposed an XGBoost learning method that considers risk potential fields, using eye movements, head movements, and Electroencephalography (EEG) to predict the quality of takeovers in conditionally automated driving under different levels of cognitive load.

However, the machine learning models are black boxes and cannot inform why and how certain factors can influence takeover performance. Thus, researchers also explored the impact of driver states and scenario features on drivers' takeover performance in system-initiated takeover scenarios. For example, previous research has found that high workload can delay drivers' responses to the takeover request (TOR) (Du et al., 2020), and trust can also affect drivers' performance in takeover scenarios (Payre et al., 2016). However, the association cannot inform the causality; thus, it is still unclear why the features can be used to effectively estimate takeover performance. For example, the HR and skin conductance level (SCL) have been used for the estimation of takeover response time (Du et al., 2020). Still, it is unknown whether the HR and SCL are affected by the driver's workload (He et al., 2022; Shi et al., 2024; Wang, et al., 2024), then the latter affects takeover performance, or the HR and SCL are directly associated with the takeover performance. Without this knowledge, the selection of features for takeover performance evaluation can only be based on trial and error.

To answer this question, a driving simulator experiment was conducted, with the drivers' physiological metrics that are commonly used for takeover performance estimation collected and drivers' psychological states that were found to be associated with the takeover performance moderated. Given that linear regressions can hardly reveal the structural relationships among the variables, we adopted a structural equation model (SEM) to explore the inter-relationships among the variables of interest. Traditionally, the SEM has been widely adopted for the analysis of questionnaire data (Chen & Donmez, 2016; Wang et al., 2022). In recent years, research has started to use SEM to explore the structures among empirical data, for example, Jin et al. (2021) used eye-tracking data as the measurement variables of latent variable. Thus, following these previous approaches, our study aims to explore the metrics potentially associated with takeover performance. The outcome of this research can guide the selection of features for takeover performance prediction, which serves as the basis for adaptive driving automation design.

2 Literature Review and Hypotheses Formation

We developed a theoretical framework for the model by formulating reasonable hypotheses (H1-H7) based on previous literature.

2.1 Physiological Signals, Eye-tracking Behavior, Workload, and Takeover Performance

Physiological signals and eye-tracking metrics have been widely recognized as being associated with takeover performance. For example, Du et al. (2020) found that during the takeover, shorter TOR lead time led to a lower frequency of blinks and higher peak and average GSR activations. Du et al. (2020) also found that a combination of HR, GSR, and eye-tracking bahavior (ETB, including fixation, saccade, pupil dilation, and blink) can reliably predict takeover performance. In another study, Zhu et al. (2023) also found that the ET, EEG signals, and head movement indicators can be used to predict takeover quality. Therefore, we proposed our first hypothesis:

H1: Physiological signals and ETB can predict takeover performance, including takeover time and takeover quality.

Further, empirical studies have demonstrated that changes in respiration (RESP) patterns (Muth et al., 2012), electrodermal activity (EDA) (Mehler et al., 2012; Widyanti et al., 2017) , and electrocardiography (ECG), as indicated by HR and HR variability (HRV) (He et al., 2019; Wang et al., 2024) are associated with workload. Similarly, ETB, including blink rate and saccadic movements, have been used to assess workload during automated driving (Aygun et al., 2022; Bitkina et al., 2021; Das & Maiti, 2024; He et al., 2022). Besides, high workload levels are known to impair drivers' ability to respond to TOR promptly, leading to increased takeover time and reduced takeover quality (Liu et al., 2024; Yoon & Ji, 2019). Therefore, we proposed the following two hypotheses:

H2: Physiological signals and ETB metrics can predict drivers' workload levels.

H3: Workload level can affect takeover performance.

2.2 Physiological States, Eye-tracking Behavior, Trust, and Takeover Performance

Trust in automation is defined as the belief in the ability of a system to perform driving tasks reliably, safely, and as expected in various conditions (Hoff & Bashir, 2015). Similar to its role in interpersonal relationships, trust in technology plays a dominant role in determining one's willingness to rely on automated systems under uncertainty (Zeeb et al., 2017) and use new technologies (Wang et al., 2021; Wang et al., 2024). Increased trust in automation, as measured by questionnaires has been found to be associated with longer takeover times (Payre et al., 2016) and shorter minimum time-to-collision (TTC) (Körber et al., 2018). Previous research also tried to measure trust objectively (e.g., using eye-tracking metrics, Körber et al., 2018). For example, HR (Khalid et al., 2016) and GSR (Ayoub et al., 2023) have also been found to be associated with driver trust in vehicles with conditional automation and Yi et al. (2023) used HR and GSR to predict drivers' trust in automation before and after takeover events. Thus, we propose the following hypotheses:

H4: Physiological signals and ETB metrics can predict trust in driving automation.

H5: Trust in driving automation can affect takeover performance.

2.3 Correlations among Physiological Signals and ETB

Studies have shown that HR is negatively associated with the rate of breath, and lower breath rates generally produce larger HRV amplitude compared to higher breath rates (Song & Lehrer, 2003). At the same time, changes in respiratory rate can lead to variations in EDA, which is particularly pronounced when there are emotional fluctuations (Li & Liu, 2011). Further, ET metrics, particularly pupil diameter, are influenced by the autonomic nervous system, and sympathetic activation can affect both pupil dilation and HR/HRV (Chen & Epps, 2014). Therefore, we proposed the following hypothesis:

H6: Interrelationships among physiological and ET metrics exist.

2.4 Correlation between Takeover Time and Quality

Previous studies have found that a longer takeover time can be associated with worse takeover quality in terms of maximum lateral acceleration (Liu et al., 2024). Further, takeover reaction time primarily depends on the length of the takeover lead time. Within a certain range, a longer lead time results in a safer takeover; however, the increased workload caused by longer lead times is associated with shorter reaction times and lower safety (Wu et al., 2022). Therefore, we proposed the following hypothesis:

H7: Takeover time affects takeover quality.

2.5 Correlation between Exogenous Variables and Latent Variables

Given the individual heterogeneity (IH) in physiological signals, psychological states, and takeover performance (Mehler et al., 2008), we further hypothesized that IH has a significant impact on these factors. Specifically, the IH variable was materialized as a unique identifier of each participant. Additionally, since takeover performance may vary across different takeover scenarios (Du et al., 2024), we proposed that takeover scenarios can influence both takeover quality and takeover time. Given that we aim to evaluate the variables that can potentially be used for takeover performance prediction, the physiological, eye-tracking and psychological variables were extracted from the period before TORs and thus are unaffected by the scenarios.

2.6 Framework of Structural Equation Model

Figure 1 summarizes the theoretical framework based on the hypotheses mentioned above. The paths from explicit variables (i.e., IH and takeover scenarios) to latent variables (i.e., RESP, EDA, ETB, ECG, TRU, WL, TOQ and TOT) are not included in the structural model diagram because these variables often serve as covariates variables influencing the latent constructs indirectly, and are typically accounted for outside the hypothesized structural relationships among the latent variables.

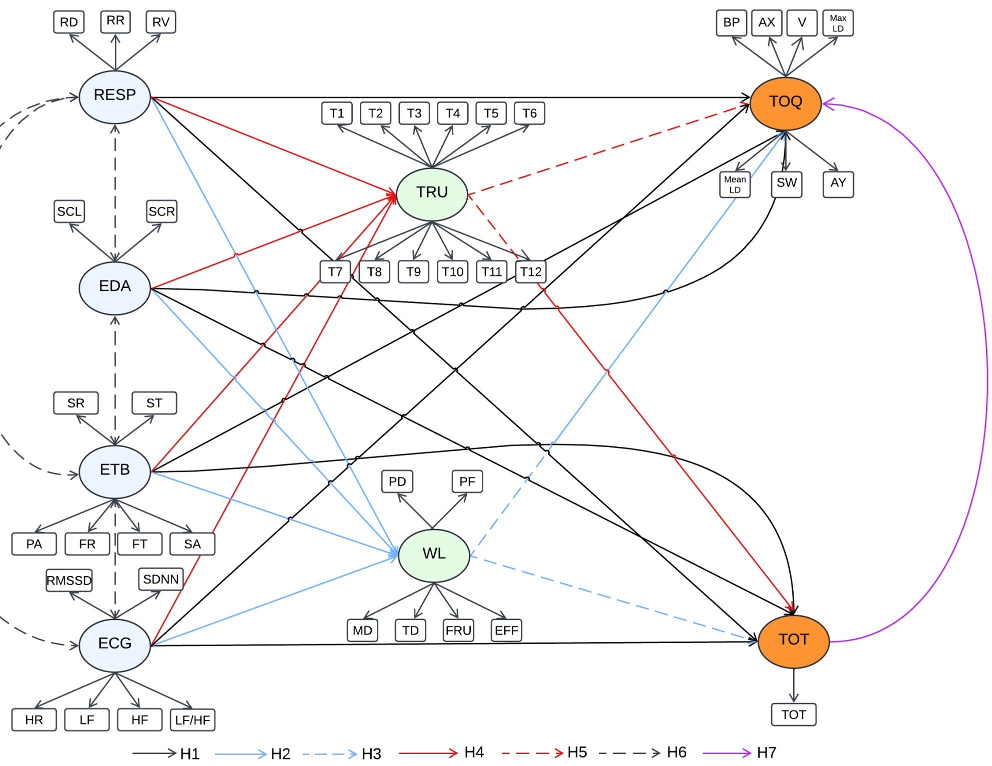

Figure 1. The final theoretical framework of the model. RESP: respiration; EDA electrodermal activity; ECG: electrocardiography; ETB: eye-tracking behavior; TRU: trust; WL: Workload; TOQ: takeover quality; and TOT: takeover time. The abbreviations in the figure are listed in Table 2.

3 Experiment

3.1 Participants

In total, 42 drivers participated in our experiment (25 males and 17 females), with an average age of 35.3 years (Standard deviation [SD]: 9.10, min: 23, max: 53). All participants were required to have a valid driver's license for a minimum of one year, no advanced driving assistant system (ADAS) usage experience and a minimum driving mileage of 5,000 km in the past year. They were told to be compensated at 70 RMB per hour plus a performance-based bonus of up to 30 RMB for cognitive tasks provided in the experiment. This study was approved by the Hong Kong University of Science and Technology (HREP-2023-0199).

3.2 Equipment

The experiment was conducted in a fixed-base driving simulator (Figure 2a), which has three 43-inch displays, showing a horizontal view angle of 150° and a vertical viewing angle of 47°. Participants could engage the driving automation by pressing the virtual buttons on the 15-inch screen next to the steering wheel. To disengage the automation, they could press another virtual button, turn the steering wheel, or press the braking pedal. The driving data was logged at a frequency of 60 Hz by the simulation software, i.e., the Silab 7.1 by WIVW GmbH. The ECG, RESP, and EDA data were collected using the sensors by Ergoneers GmbH at a sampling frequency of 100 Hz (see Figure 2b). The ET measures were collected using the Dikablis and D-Lab software with a sampling frequency of 60 Hz. All data was synchronized in the Human Research Tool (HRT) software by Info Instrument.

(a) (b)

Figure 2. Equipment: (a) driving simulator and eye-tracker; b) physiological sensor placements.

3.3 Driving Task

Each participant completed one drive on a dual carriageway with six lanes, which has 21 straight sections, each stretching around 7 km. The speed limit was 110 km/h and the traffic was free-flowing, with a vehicle density of 6 passenger cars per km per lane. All participants were informed that the vehicle was equipped with SAE Level 3 driving. Thus, they were required to engage the driving automation when possible and only take over the control of the vehicle when they felt necessary or prompted by TOR. In other words, if a TOR was triggered, they should take over the control of the vehicle. Nevertheless, it turns out that, in the experiment, no participants took control of the vehicle before TOR and all takeover actions were triggered by the TOR. As shown in Figure 3, three types of takeover events were used in the drive, i.e., exiting the ramp, traffic accidents, and foggy areas. In the scenario of exiting the ramp, drivers must change lanes to the right lane, then enter the right-hand exit ramp, and control the speed to be below 60 km/h while on the ramp. In the scenario of a traffic accident, after taking over the vehicle, drivers needed to change lanes to avoid an accident area. In the scenario of foggy areas, drivers only needed to reduce speed and proceed through the fog zone after taking over.

Before each takeover event, a TOR was provided, with a lead time (from the TOR initiation to automation disengagement) of 10 seconds following previous research (Chen et al., 2023).

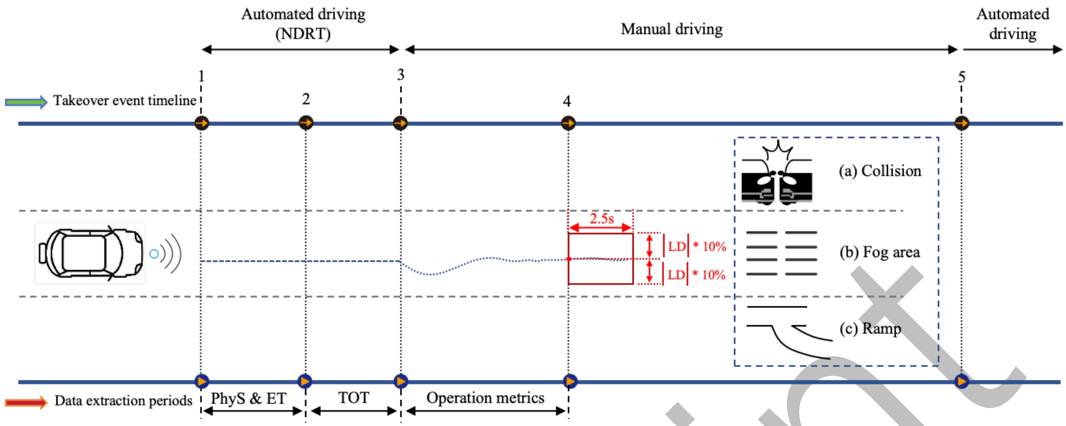

Figure 3. Scenario design. In the takeover event timeline, 1, 2, 3, 4, and 5 marks the moments of event start, TOR, ADAS disengagement, vehicle operation stabilization and the end of the event, respectively. Further, NDRT is *Non-Driving Related Task*, TOR is *Takeover Request*, TOT is *Takeover Time*, LD is *Lateral Deviation*, PhyS is *Physiological Signal*, and ET is *Eye-Tracking*. The red section indicates the point at which the vehicle achieves lateral stability after the takeover, defined as the moment when the LD remained within 10% of the maximum lateral deviation after ADAS disengagement for at least 2.5 seconds (Chen et al., 2023).

3.4 Non-Driving-Related Tasks (NDRTs)

The workload is not unidimensional and may occupy different types of cognitive resources (Wang et al., 2024). To evaluate how each dimension can affect takeover performance, as shown in Table 1 and Figure 4, three types of cognitive tasks were adopted for this study, i.e., calculation, memory, and spatial processing, which are directly related to the cognitive functions for safe driving. Specifically, calculation can be used for estimating the remaining range of electric vehicles, memory is involved in processing and recalling traffic scenarios, and spatial processing is used for navigation. We further controlled the level of distraction tasks to impose different levels of workload, simulating different traffic complexities. As a result, we have in total seven NDRT conditions (3 memory task levels, 2 calculation task levels, one spatial processing task and the baseline without NDRT) in the experiment.

Table 1. Summary of NDRTs.

Task	Descriptions	Levels	Cognitive Resource
n-back Task (Jaeggi et al., 2010)	A series of stimuli (numbers/letters) are presented with a pause between each. Participants say out the stimulus that is *n* positions earlier.	**Three levels**: 0-back (0-B), 1-back (1-B), 2-back (2-B) tasks.	Memory
Math Task (Meteier et al., 2021)	Oral backward counting from 3,000 by increments of 3 (non-integer) or 5 (integer).	**Two levels**: Counting backward by 3 (MT1) or 5 (MT2) from 3,000.	Calculation
Spatial Task (Liang & Lee, 2010)	Participants listen to an audio clip describing a route and say out the main direction faced at the end, simulating high cognitive demands similar to those in navigation systems.	**One level**: "What direction is this person facing when she goes to the north station and moves two stations clockwise?" (Answer: East) (ST)	Spatial processing

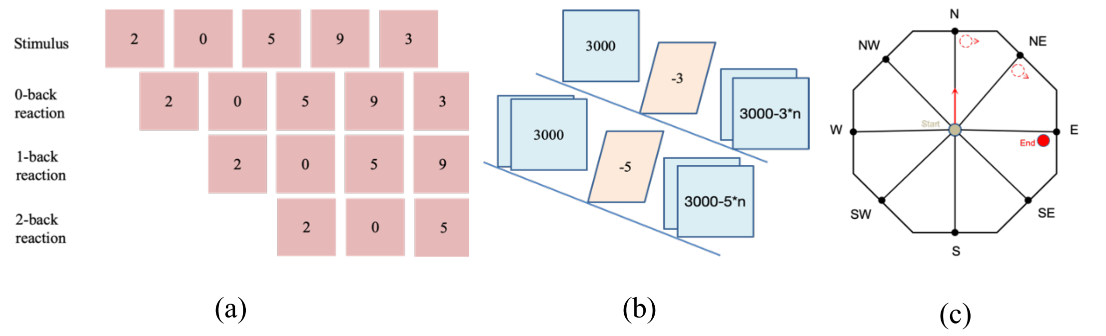

Figure 4. Cognitive tasks: (a) n-back task, (b) math task, and (c) spatial task.

It turns out that the accuracies of the responses to the six tasks were 98.5% (SD: 2.4%), 85.2% (SD: 8.0%), 76.3 (SD: 8.0%), 79.1% (SD: 5.9%), 82.67% (5.3%) and 80.1% (SD: 9.7%) for the 0-back, 1-back, 2-back, MT1, MT2 and ST tasks, respectively. These accuracies were comparable to previous studies that used the n-back tasks (Miller et al., 2009), indicating that our participants were fully engaged in these tasks.

3.5 Experiment Design

A within-subject experiment design was adopted. The experiment employed a 3 (takeover scenarios) by 7 (NDRT conditions) within-subject design, leading to 21 drives for all combinations of experiment conditions. The takeover scenario happened near the end of each 3-minute drive. We used a Latin square design to balance the sequence and reduce learning effects, leading to 21 unique drives and 21 experimental orders. Each order involved two participants, who were required to complete all experimental conditions. After completing each 7 drives, participants took a 10-minute break before proceeding to the next 7 drives.

3.6 Procedures

Participants were reminded to maintain regular sleep habits, avoid alcohol, and not consume caffeine 24 hours before the experiment. As shown in Figure 5, upon arrival, participants provided written consent, followed by a 30-minute orientation session regarding the experimental procedures, operation of the vehicles, and cognitive tasks. Then, participants underwent a practice driving session, in which the driver experienced one TOR in on highway. Afterward, the participants were equipped with physiological sensors and eye-tracking devices, and then the formal experiment began, in which, each participant experienced 21 TORs, i.e., 3 takeover scenarios in combination with 7 task load levels (baseline, 3 n-back levels, 2 math task levels, and 1 cognitive spatial task). Each TOR happened on one of the 21 straight sections of the road. The order of the experiment conditions was counterbalanced using a Latin-Square design. The driver was required to complete two questionnaires, i.e., the NASA Task Load Index (NASA_TLX) (Hart & Staveland, 1988) for workload, and a 12-item questionnaire by Jian et al. (2000) for trust in driving automation. To avoid potential impacts of takeover events on subjective data reporting, we reminded participants during the training phase and each time before the questionnaire that they should base their responses on the duration before the takeover events.

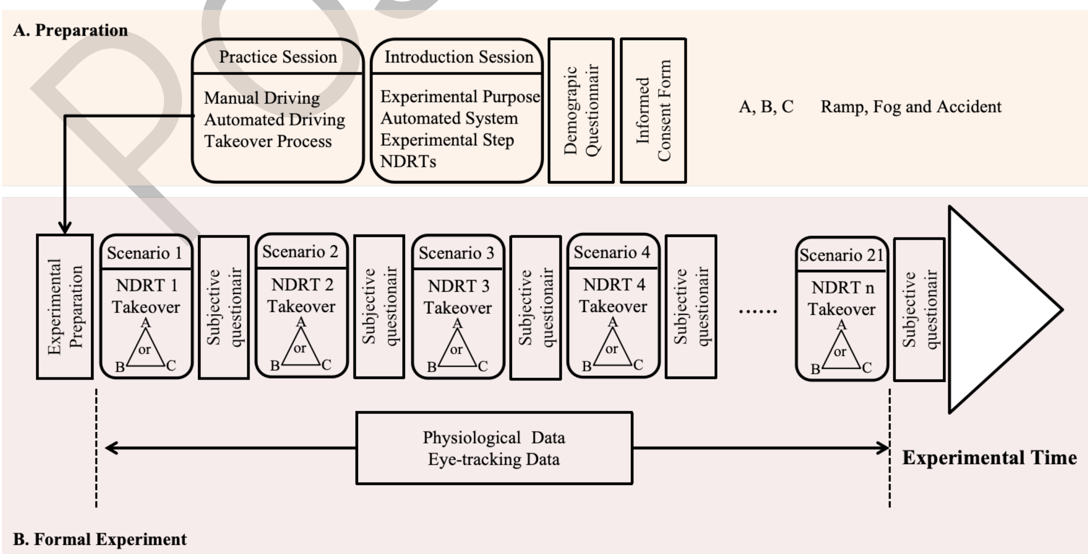

Figure 5. Experimental process.

3.7 Signal Processing and Variable Extraction

As shown in Table 2 and Figure 3, the variables of interest regarding drivers' takeover performance, physiological responses, and psychological states were extracted. The takeover time was from the moment the TOR was initiated to the deactivation of the ADAS. The time period for extracting takeover quality data was from the ADAS disengagement to the achievement of lateral stability. It should be noted that for the physiological measures, band-pass filters were applied to reduce the noise in the data before metric extraction. Specifically, for EDA, a fourth-order Butterworth low-pass filter was applied to mitigate high-frequency disturbances. ECG signals were processed through a band-pass filter between 3 Hz to 45 Hz, followed by R-wave detection using an enhanced Pan-Tompkins algorithm (Sathyapriya et al., 2014) to calculate the R-R intervals. For RESP, band-pass filters were applied within the frequency range between 0.1 Hz to 0.35 Hz. All physiological metrics and the ET metrics were extracted 120 seconds before TOR to the TOR onset; while the psychological states were based on the corresponding questionnaire responses collected after each takeover event.

Table 2. Extraction of variables.

Measures	Metrics	Unit	Definitions	Calculation Details
Workload (WL)	MD	-	Mental load	Each dimension served as a measurement item for the latent variable WL in SEM, each ranging from 1 (low) to 21 (high). (Hart & Staveland, 1988)
	PD	-	Physical load	
	TD	-	Temporal load	
	EFF	-	Effort degree	
	PF	-	Performance satisfaction	
	FRU	-	Frustration degree	
Trust (TRU)	T1	-	Deceptive	Seven-point Likert scale, ranging from 1 "strongly disagree" to 7 "strongly agree." (Jian et al., 2000). Each dimension served as a measurement item for the latent variable TRU in the SEM.
	T2	-	Underhanded	
	T3	-	Suspicious	
	T4	-	Wary	
	T5	-	Harmful or injurious	
	T6	-	Confident	
	T7	-	Security	
	T8	-	Integrity	
	T9	-	Dependable	
	T10	-	Reliable	
	T11	-	Trust	
	T12	-	Familiar	
Takeover time	TOT	s	Takeover time, i.e., The duration from the issuance of a takeover request	
Takeover quality	V	m/s	Vehicle speed	The average/max value of the
	Max LD	m	The max lateral distance from the centerline of the lane	

(TOQ)	Mean LD	m	The mean lateral distance from the centerline of the lane	metrics during the period from the initiation of takeover to effective takeover (lateral stability). (Cao et al., 2021; Yao et al., 2021).
	BP	-	Vehicle brake pedal depth (0-10)	
	SW	rad	Vehicle steering wheel angel	
	AX	m/s^2	Longitudinal acceleration	
	AY	m/s^2	Lateral acceleration	
Eye-tracking (ET) metrics	PA	px	Pupil area	Average of the metrics from 120 seconds prior to the TOR to the moment of the TOR initiation.
	FR	Times/minute	Fixations rate, i.e., the number of fixations per minute, where a fixation	
	FT	s/min	Fixations time, i.e., the total fixation duration divided by the total time.	
	SR	Times/minute	Saccade rate, i.e., the number of saccade per minute.	
	ST	s/min	Saccade time, i.e., the total saccade duration divided by the total time	
	SA	Degrees	Saccade angle, i.e., the angle between two consecutive fixation points	
Electrodermal Activity (EDA)	SCL	µS	Skin conductance level	
	SCR	µS	Skin conductance response	
Electrocardiogram (ECG)	HR	Beats/minute	Heart rate	
	RMSSD	s	Magnitude of the differences between consecutive R-R intervals	
	SDNN	s	Variability of the time intervals between consecutive normal heartbeats	
	LF	ms^2	Spectral power in the low-frequency range (usually 0.04 to 0.15 Hz)	
	HF	ms^2	Spectral power in the high-frequency range (usually 0.15 to 0.4 Hz)	
	LF/HF	%	Ratio of LF to HF	
Respiration (RESP)	RR	Times/ minute	Rate of breaths	
	RD	mm	Size of respiratory depth	
	RV	%	Variation in respiratory intervals	

3.8 Construction of Structural Equation Model

A three-step approach was employed to test the research hypotheses in SEM models, following the methodology by Okumus et al. (2016). First, the measurement models were constructed to evaluate the validity of the model, by analyzing the relationships between measurement items and latent variables through a confirmatory factor analysis. At the same time, the reliability of the model was assessed through composite reliability (CR), and the convergent validity was evaluated by calculating the average variance extracted (AVE). Next, SEM analysis was performed to calculate the standardized coefficients for each factor. The goodness-of-fit of the model was evaluated using the criteria recommended by Schermelleh-Engel et al. (2003). Specifically, the Chi-square/Degree of Freedom (χ^2/df) index measures the overall fit by comparing the proposed model structure to the observed data. The Goodness of Fit Index (GFI) reflects how well the model explains the variance in the observed data. The Adjusted Goodness-of-Fit Index (AGFI) further adjusts for model complexity and degrees of freedom. The Normed Fit Index (NFI) evaluates the fit by comparing the proposed model to a null model. The Comparative Fit Index (CFI) also compares the model to a baseline, taking sample size into account. Finally, the Root Mean Square Error of Approximation (RMSEA) assesses the model by measuring the error of approximation per degree of freedom.

4 Results

4.1 Validation of the Cognitive Tasks

We first checked if the NDRT we used imposed differentiable loads on drivers. It was found that the NASA_TLX data did not follow a normal distribution, so we used the Friedman test to compare workload levels across varying task difficulties within each type of cognitive task. When a significant effect was detected, we conducted Wilcoxon signed-rank tests for post-hoc comparisons. As illustrated in Figure 6, our task imposed distinct levels of workload on drivers.

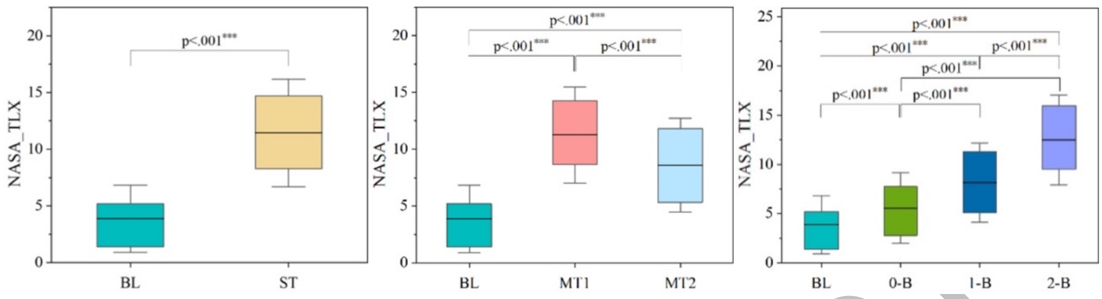

Figure 6. Significance test results of self-reported NASA-TLX. (a) Cognitive spatial task, (b) Math task, and (c) N-back task. Note: MT1 is math task 1 (decrement of 3), MT2 is math task 2 (decrement of 5), BL is the baseline (driving only), 0-B is 0-back task, 1-B is 1-back task and 2-B is the 2-back task, *** means $p < .001$.

4.2 SEM

The SEM constructed 25 paths based on 812 valid data points (42 participants and 21 takeovers per participant, with 70 data points being lost due to physiological or eye-tracking data missing). To account for the correlations introduced by repeated measurements (as each participant experienced multiple takeover events), we adopted the Mixed-Effects Model (MLM) in SEM to account for individual and experimental sequence effects. Thus, the model can handle the repeated measures on IH. This approach avoids the violation of the SEM assumptions on data independence.

Table 3 presents the results of the measurement model. For the measurement model, measurement items with factor loadings below 0.5 were removed, including RR from RESP, HR and LF/HF from ECG, ST, FT and SA from ETB, PD and PF from WL, Familiarity from TRU, and V, Mean LD, AY, and SW from TOQ. As a result, we found that the SCL, RD, SDNN, FR, MD, T9 of Trust, and BP were the most effective indicators of EDA, RESP, ECG, ETB, WL, TRU, and TOQ, respectively, as indicated by their high loading factors. At the same time, the CR of all constructs ranged from 0.768 to 0.95, all above the recommended threshold of 0.7 (Wang et al., 2021), indicating internal consistency. The AVE values all ranged from 0.513 to 0.748, surpassing the recommended threshold of 0.5 (Wu et al., 2023), indicating the convergent validity of the model.

Table 3. Results of factor analysis.

Latent variable	Measurement Variable	Factor loading	AVE	CR
EDA	SCL	0.824	0.624	0.768
	SCR	0.755		
RESP	RD	0.903	0.684	0.810
	RV	0.743		
ECG	RMSSD	0.676	0.513	0.807
	SDNN	0.823		
	LF	0.640		
	HF	0.715		
ETB	SR	0.872	0.748	0.898
	PA	0.743		
	FR	0.966		
WL	MD	0.918	0.693	0.899
	TD	0.842		
	FRU	0.664		
	EFF	0.882		
TRU	T1	0.703	0.636	0.950
	T2	0.679		
	T3	0.772		
	T4	0.671		
	T5	0.701		
	T6	0.855		
	T7	0.808		
	T8	0.662		
	T9	0.973		
	T10	0.970		

	T11	0.893		
TOQ	BP	-0.839	0.583	0.807
	AX	-0.724		
	Max LD	-0.722		

As shown in Table 4, all fit indices are acceptable, indicating satisfactory goodness-of-fit of the model. Then, Figure 7 and Table 5 present the path coefficients and their significance in the final model with non-significant paths eliminated.

The results indicate that none of the physiological signals were direct predictors of TOQ (Takeover Quality) or TOT (Takeover Time). Instead, they were associated with TOQ and TOT indirectly through psychological latent variables (WL, TRU). Specifically, ETB (Eye-tracking behavior) and EDA (Electrodermal Activity) were related to TOQ and TOT through the mediating psychological variable TRU (Trust), with increased eye-related activity and decreased EDA activity being associated with higher levels of TRU, and consequently, higher TOQ and TOT. At the same time, ETB, ECG (Electrocardiogram) and RESP (Respiration) were related to TOT through the mediating psychological state WL, with decreased ECG, increased RESP and ETB activity being associated with higher workload, and consequently higher TOT. Additionally, we observed an association between RESP and ECG, with higher RESP activity leading to higher ECG activity. However, the ETB was not associated with any physiological signals. Finally, we found that individual heterogeneity (IH) had an impact on physiological signals, eye movement metrics, psychological states, and takeover performance, with the largest effect observed on EDA and the smallest effect on TOQ, and the takeover scenarios can affect TOQ and TOT.

Table 4. Goodness-of-fit of the SEM model.

Goodness-of-fit measures	Good fit	Acceptable fit	Model value
Chi-square/Degree of Freedom (χ^2/df)	[0,2]	(2,3)	2.574
Goodness of Fit Index (GFI)	[0.95,1]	[0.90,0.95)	0.973
Adjusted Goodness-of-Fit Index (AGFI)	[0.90,1]	[0.85,0.90)	0.966
Normed Fit Index (NFI)	[0.95,1]	[0.90,0.95)	0.952
Comparative Fit Index (CFI)	[0.97,1]	[0.95,0.97)	0.931
Root Mean Square Error of Approximation (RMSEA)	[0,0.05]	(0.05,0.08]	0.059

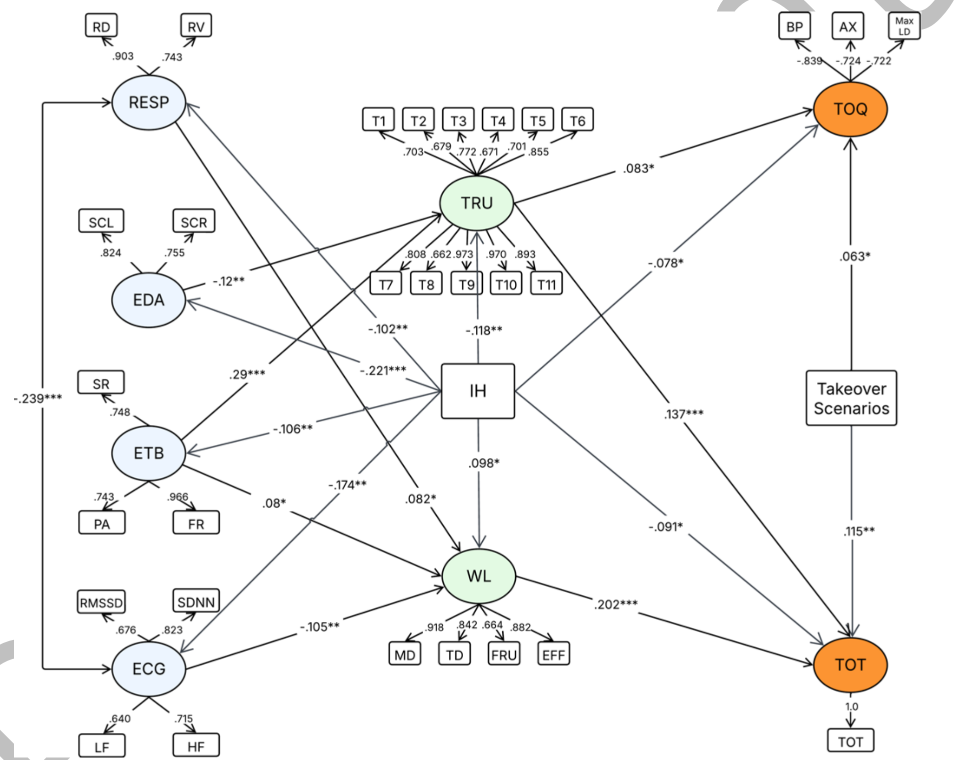

Figure 7. The final model results. * means $p < .05$, ** means $p < .01$, *** means $p < .001$.

Only significant paths are shown.

Table 5. Mediating effect analysis.

Physiological	Takeover	Mediating	Effect path	Effect size
RESP	TOT	WL	Total	0.018
			RESP→WL	0.082
			TRU→TOQ	0.225
EDA	TOQ	TRU	Total	-0.01
			EDA→TRU	-0.12

			TRU→TOQ	0.083
	TOT	TRU	Total	-0.02
			EDA→WL	-0.12
			WL →TOT	0.137
ECG	TOT	WL	Total	-0.02
			ECG→WL	-0.101
			WL →TOT	0.202
ETB	TOT	WL and TRU	Total	0.056
			ETB→WL	0.08
			WL →TOT	0.202
			ETB→TRU	0.29
			TRU →TOT	0.137
	TOQ	TRU	Total	0.024
			ETB→TRU	0.29
			TRU →TOQ	0.083

5 Discussion

Based on the data from a driving simulator study, we evaluated the relationships among users' psychological states, physiological states, eye-tracking metrics, and takeover performance in SAE Level-3 vehicles. In our model, the smaller TOQ value was associated with larger brake pedal depth (BP), longitudinal acceleration (AX), and maximum lane deviation (MaxLD), which all indicate worse takeover quality (Cao et al., 2021; Jin et al., 2021; Yao et al., 2021). Thus, the better the takeover quality, the larger the TOQ value. Similarly, the larger the TOT value, the longer the takeover time, as the TOT is only positively associated with the takeover time.

First, in line with previous research (Du et al., 2020; Liu et al., 2024; Payre et al., 2016) and partially supporting H3 and H5, we found that the driver's trust in driving automation was associated with both takeover quality and takeover time. In contrast, workload was only associated with takeover time. Specifically, we first found that trust was positively associated with takeover quality but not takeover time. The positive association between trust and takeover quality indicates that drivers would generate smoother and stabler control of the vehicle when they trusted more in driving automation. It is possible that those who pose higher trust in driving automation would be less stressed when taking back control of the vehicle, which can partially be supported by the negative relationship between the EDA and the trust in the SEM model (supporting H4), as previous research has found that EDA was associated with increased stress among drivers (Setz et al., 2010). Further, in line with Jin et al. (2021), we also observed a positive association between trust and takeover time. This is as expected, as drivers who trusted in automation more might tend to rely on automation more and thus respond more slowly.

Conversely, partially supporting H3, workload was associated with takeover time, but not takeover quality. The significant positive relationship between workload and takeover time agrees with the findings in previous driving studies (Liu et al., 2024), as switching between ongoing (i.e., NDRT) and interrupting tasks (i.e., driving task) is cognitively demanding and thus time-consuming (Payre et al., 2016). The lack of association between workload and takeover quality is also reasonable. In the takeover process, drivers would relocate their cognitive resources from NDRT to the driving task (Zeeb et al., 2016). Thus, the takeover performance would not be compromised if the shift of attention can be completed in time. However, future research is still needed to further validate our explanations, potentially through more direct measures of drivers' brain activities, such as electroencephalogram (EEG) or Functional Near-Infrared Spectroscopy (fNIRS).

Contradictory to our expectation (H1), physiological activity was not directly associated with the takeover performance but indirectly correlated with the takeover performance through trust and workload. This hierarchical relationship suggests that when using physiological signals to predict takeover performance (e.g., Du et al., 2020), the machine learning models might be estimating the psychological states of the drivers. Hence, additional information that can explain the variations in drivers' psychological states, such as the traffic conditions that affect drivers' task load (Stapel et al., 2019) and environmental factors that influence situational trust (Jin et al., 2021), may improve the accuracy of takeover performance estimation.

Additionally, partially supporting H6, we observed a significant correlation between respiratory activity and electrocardiographic signals, suggesting the underlying relationships or mechanisms among different physiological measures (Wang et al., 2024). Further, partially supporting H2 to H5, several physiological metrics and eye-tracking metrics were indirectly related to takeover time and takeover quality through psychological states. However, it should be noted that being different from previous studies (He et al., 2019), the EDA was not found to be associated with the workload. It is likely that the variations in other drivers' states shadowed the variation in EDA, given that the EDA is sensitive to multiple states of the drivers (e.g., Radhakrishnan et al., 2022; Sarchiapone et al., 2018).

Next, being contradictory to H7, takeover time was not associated with takeover quality, which suggests that a rapid response (short TOT) does not necessarily ensure a high-quality takeover. It is likely that the takeover time was more related to the takeover scenario and drivers' capability to rebuild situation awareness when the vehicle was controlled by ADAS (Tanshi & Söffker, 2019), while the takeover quality was more related to drivers' manual driving skills (Soares et al., 2021). Finally, we found that all variables of interest in our study were associated with individual heterogeneity to some extent, suggesting the need to consider individual differences when estimating drivers' performance in takeover events.

6 Limitations

This study has several limitations that should be considered when interpreting the findings. First, the realism of the driving simulator might influence the participants' physiological responses. Specifically, while driving simulators provide a safe and controllable environment for experiments, they may not fully replicate the complexity of real-world driving, which could potentially bias biometric signals. Future studies should validate these findings in both simulator and on-road experiments to assess their generalizability (such as Yang et al., 2025). Secondly, only cognitive tasks were used in the study. Further research should be conducted to validate the conclusions when NDRTs with other modalities (e.g., visual and manual) are provided.

7 Conclusion

In this study, we adopted a structural equation model to explore the relationships among physiological signals (i.e., cardiac activity, respiratory activity, electrodermal activity), eye-tracking metrics, psychological states (i.e., trust, workload), and variations in takeover performance (takeover time and takeover quality) in conditionally automated vehicles. The major findings are summarized as follows:

- The driver's psychological states (i.e., trust in the automated system and workload) were directly associated with takeover performance. Specifically, trust was positively associated with takeover quality and takeover time, whereas drivers' task load was positively associated with takeover time.
- Physiological signals (including respiratory activity, cardiac activity, and electrodermal

activity) and eye-tracking metrics were found to be indirectly associated with takeover performance via drivers' psychological states, but they were not directly associated with takeover performance.

- Physiological signals and eye-tracking metrics are associated with psychological states.

This study reveals the hierarchical relationship among physiological signals, eye-tracking behavior, psychological states, and takeover performance, emphasizing the influence of drivers' states on takeover performance. Understanding these complex relationships can guide the design of driver monitoring systems for takeover performance estimation.

CRediT authorship contribution statement

Ange WANG: Methodology, Software, Validation, Formal analysis, Writing – original draft. Jiyao WANG: Writing – review & editing. Chunxi Huang: Conceptualization, Writing – review & editing. Dengbo HE: Methodology, Writing – review & editing, Funding acquisition, Supervision. Hai YANG: Writing – review & editing.

Declaration of Competing Interest

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

Acknowledgments

This work was supported by the Natural Science Foundation of Guangdong Province of China (2024A1515010392), and partially by the National Natural Science Foundation of China (grant no. 52202425), the Guangzhou Municipal Science and Technology Project (No. 2023A03J0011) and Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No. 2023B1212010007).

References

Aygun, A., Lyu, B., Nguyen, T., Haga, Z., Aeron, S., & Scheutz, M. (2022). *Cognitive Workload Assessment via Eye Gaze and EEG in an Interactive Multi-Modal Driving Task*. 337–348. Scopus. https://doi.org/10.1145/3536221.3556610

Ayoub, J., Avetisian, L., Yang, X. J., & Zhou, F. (2023). Real-Time Trust Prediction in Conditionally Automated Driving Using Physiological Measures. *IEEE Transactions on Intelligent Transportation Systems*, *24*(12), 14642–14650. IEEE Transactions on Intelligent Transportation Systems. https://doi.org/10.1109/TITS.2023.3295783

Ayoub, J., Du, N., Yang, X. J., & Zhou, F. (2022). Predicting Driver Takeover Time in Conditionally Automated Driving. *IEEE Transactions on Intelligent Transportation Systems*, *23*(7), 9580–9589. https://doi.org/10.1109/TITS.2022.3154329

Bitkina, O. V., Park, J., & Kim, H. K. (2021). *The ability of eye-tracking metrics to classify and predict the perceived driving workload*. *86*. Scopus. https://doi.org/10.1016/j.ergon.2021.103193

Cao, Y., Zhou, F., Pulver, E. M., Molnar, L. J., Robert, L. P., Tilbury, D. M., & Yang, X. J. (2021). Towards standardized metrics for measuring takeover performance in conditionally automated driving: A systematic review. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *65*(1), 1065–1069. https://doi.org/10.1177/1071181321651213

Chen, H., Zhao, X., Li, Z., Fu, Q., Wang, Q., & Zhao, L. (2023). Construction and Analysis of Driver Takeover Behavior Modes Based on Situation Awareness Theory. *IEEE Transactions on Intelligent Vehicles*, 1–15. https://doi.org/10.1109/TIV.2023.3280077

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

Chen, H.-Y. W., & Donmez, B. (2016). What drives technology-based distractions? A structural equation model on social-psychological factors of technology-based driver distraction engagement. *Accident Analysis & Prevention*, *91*, 166–174. https://doi.org/10.1016/j.aap.2015.08.015

Chen, S., & Epps, J. (2014). Using Task-Induced Pupil Diameter and Blink Rate to Infer Cognitive Load. *Human–Computer Interaction*. https://www.tandfonline.com/doi/abs/10.1080/07370024.2014.892428

Das, S., & Maiti, J. (2024). Assessment of cognitive workload based on information theory enabled eye metrics. *Safety Science*, *176*, 106567. https://doi.org/10.1016/j.ssci.2024.106567

Du, N., Kim, J., Zhou, F., Pulver, E., Tilbury, D. M., Robert, L. P., Pradhan, A. K., & Yang, X. J. (2020). Evaluating Effects of Cognitive Load, Takeover Request Lead Time, and Traffic Density on Drivers' Takeover Performance in Conditionally Automated Driving. *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 66–73. https://doi.org/10.1145/3409120.3410666

Du, N., Yang, X. J., & Zhou, F. (2020). Psychophysiological responses to takeover requests in conditionally automated driving. *Accident Analysis & Prevention*, *148*, 105804. https://doi.org/10.1016/j.aap.2020.105804

Du, N., Zhou, F., Pulver, E. M., Tilbury, D. M., Robert, L. P., Pradhan, A. K., & Yang, X. J. (2020). Predicting driver takeover performance in conditionally automated driving. *Accident Analysis and Prevention*, *148*. Scopus. https://doi.org/10.1016/j.aap.2020.105748

Du, N., Zhou, F., Tilbury, D. M., Robert, L. P., & Jessie Yang, X. (2024). Behavioral and physiological responses to takeovers in different scenarios during conditionally automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *101*, 320–331. https://doi.org/10.1016/j.trf.2024.01.008

Hao LI & Jie LIU. (2011). Measurement and analysis of the relationship between human skin resistance and respiration and emotion. *Science Technology and Engineering*, *12*(3), 1671–1815.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

He, D., Donmez, B., Liu, C. C., & Plataniotis, K. N. (2019). High Cognitive Load Assessment in Drivers Through Wireless Electroencephalography and the Validation of a Modified N-Back Task. *IEEE Transactions on Human-Machine Systems*, *49*(4), 362–371. IEEE Transactions on Human-Machine Systems. https://doi.org/10.1109/THMS.2019.2917194

He, D., Wang, Z., Khalil, E. B., Donmez, B., Qiao, G., & Kumar, S. (2022). Classification of Driver Cognitive Load: Exploring the Benefits of Fusing Eye-Tracking and Physiological Measures. *Transportation Research Record: Journal of the Transportation Research Board*, *2676*(10), 670–681. https://doi.org/10.1177/03611981221090937

Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394–412. https://doi.org/10.1080/09658211003702171

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04

Jin, M., Lu, G., Chen, F., Shi, X., Tan, H., & Zhai, J. (2021). Modeling takeover behavior in level 3 automated driving via a structural equation model: Considering the mediating role of trust. *Accident Analysis & Prevention*, *157*, 106156. https://doi.org/10.1016/j.aap.2021.106156

Khalid, H. M., Shiung, L. W., Nooralishahi, P., Rasool, Z., Helander, M. G., Kiong, L. C., & Ai-vyrn, C. (2016). Exploring Psycho-Physiological Correlates to Trust. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. https://doi.org/10.1177/1541931213601160

Körber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, *66*, 18–31. https://doi.org/10.1016/j.apergo.2017.07.006

Körber, M., Prasch, L., & Bengler, K. (2018). Why Do I Have to Drive Now? Post Hoc Explanations of Takeover Requests. In *Human Factors* (Vol. 60, Issue 3, pp. 305–323). https://doi.org/10.1177/0018720817747730

Liang, Y., & Lee, J. D. (2010). Combining cognitive and visual distraction: Less than the sum of its parts. *Accident Analysis & Prevention*, *42*(3), 881–890. https://doi.org/10.1016/j.aap.2009.05.001

Liu, Y., Wu, C., Zhang, H., Ding, N., Xiao, Y., Zhang, Q., & Tian, K. (2024). Safety evaluation and prediction of takeover performance in automated driving considering drivers' cognitive load: A driving simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *103*, 35–52. https://doi.org/10.1016/j.trf.2024.03.021

Mehler, B., Reimer, B., & Coughlin, J. F. (2012). Sensitivity of Physiological Measures for Detecting Systematic Variations in Cognitive Demand From a Working Memory Task: An On-Road Study Across Three Age Groups. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*(3), 396–412. https://doi.org/10.1177/0018720812442086

Mehler, B., Reimer, B., Pohlmeyer, A. E., & Coughlin, J. (2008). The association between heart rate reactivity and driving performance under dual task demands in late middle age drivers. *Advances in Transportation Studies*, *SPEC.ISSU*, 53–70. Scopus.

Meteier, Q., Capallera, M., Ruffieux, S., Angelini, L., Abou Khaled, O., Mugellini, E., Widmer, M., & Sonderegger, A. (2021). Classification of Drivers' Workload Using Physiological Signals in Conditional Automation. *Frontiers in Psychology*, *12*. Scopus. https://doi.org/10.3389/fpsyg.2021.596038

Miller, K. M., Price, C. C., Okun, M. S., Montijo, H., & Bowers, D. (2009). Is the N-Back Task a Valid Neuropsychological Measure for Assessing Working Memory? *Archives of Clinical Neuropsychology*, *24*(7), 711–717.

https://doi.org/10.1093/arclin/acp063

Muth, E. R., Moss, J. D., Rosopa, P. J., Salley, J. N., & Walker, A. D. (2012). Respiratory sinus arrhythmia as a measure of cognitive workload. *International Journal of Psychophysiology*, *83*(1), 96–101. https://doi.org/10.1016/j.ijpsycho.2011.10.011

Okumus, B., Bilgihan, A., & Ozturk, A. B. (2016). Factors Affecting the Acceptance of Smartphone Diet Applications. *Journal of Hospitality Marketing & Management*. https://www.tandfonline.com/doi/abs/10.1080/19368623.2016.1082454

Payre, W., Cestac, J., & Delhomme, P. (2016). Fully Automated Driving: Impact of Trust and Practice on Manual Control Recovery. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(2), 229–241. https://doi.org/10.1177/0018720815612319

Radhakrishnan, V., Merat, N., Louw, T., Gonçalves, R. C., Torrao, G., Lyu, W., Puente Guillen, P., & Lenné, M. G. (2022). Physiological indicators of driver workload during car-following scenarios and takeovers in highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *87*, 149–163. Scopus. https://doi.org/10.1016/j.trf.2022.04.002

Sarchiapone, M., Gramaglia, C., Iosue, M., Carli, V., Mandelli, L., Serretti, A., Marangon, D., & Zeppegno, P. (2018). The association between electrodermal activity (EDA), depression and suicidal behaviour: A systematic review and narrative synthesis. *BMC Psychiatry*, *18*(1), 22. https://doi.org/10.1186/s12888-017-1551-4

Sathyapriya, L., Murali, L., & Manigandan, T. (2014). Analysis and detection R-peak detection using Modified Pan-Tompkins algorithm. *2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies*, 483–487. https://doi.org/10.1109/ICACCCT.2014.7019490

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research*, *8*(2), 23–74.

Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., & Ehlert, U. (2010). Discriminating Stress From Cognitive Load Using a Wearable EDA Device. *IEEE Transactions on Information Technology in Biomedicine*, *14*(2), 410–417. IEEE Transactions on Information Technology in Biomedicine. https://doi.org/10.1109/TITB.2009.2036164

Shi, W., Wang, Z., Wang, A., & He, D. (2024). Classification of Driver Cognitive Load in Conditionally Automated Driving: Utilizing Electrocardiogram-Based Spectrogram with Lightweight Neural Network. *Transportation Research Record*. https://doi.org/10.1177/03611981241252797

Soares, S., Lobo, A., Ferreira, S., Cunha, L., & Couto, A. (2021). Takeover performance evaluation using driving simulation: A systematic review and meta-analysis. *European Transport Research Review*, *13*(1), 47.

https://doi.org/10.1186/s12544-021-00505-2

Song, H.-S., & Lehrer, P. M. (2003). The Effects of Specific Respiratory Rates on Heart Rate and Heart Rate Variability. *Applied Psychophysiology and Biofeedback*, *28*(1), 13–23. https://doi.org/10.1023/A:1022312815649

Stapel, J., Mullakkal-Babu, F. A., & Happee, R. (2019). Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 590–605. https://doi.org/10.1016/j.trf.2018.11.006

Tanshi, F., & Söffker, D. (2019). Modeling of takeover variables with respect to driver situation awareness and workload for intelligent driver assistance. *2019 IEEE Intelligent Vehicles Symposium (IV)*, 1667–1672. https://doi.org/10.1109/IVS.2019.8814045

Wang, A., Guan, H., Guo, J., Han, Y., & Bian, H. (2022). An Empirical Study on the Segmentation of Potential Users of Shared Parking Spaces considering Individual Heterogeneity. *Journal of Advanced Transportation*, *2022*.

Wang, A., Guan, H., Han, Y., & Cao, Y. (2021). Private Parking Space Sharing Intention in China: An Empirical Study Based on the MIMIC Model. *Discrete Dynamics in Nature and Society*, *2021*.

Wang, A., Guan, H., Qin, Z., & Zhu, J. (2021). Study on the Intention of Private Parking Space Owners of Different Levels of Cities to Participate in Shared Parking in China. *Discrete Dynamics in Nature and Society*, *2021*, e9955686. https://doi.org/10.1155/2021/9955686

Wang, A., Huang, C., Wang, J., & He, D. (2024). The association between physiological and eye-tracking metrics and cognitive load in drivers: A *meta*-analysis. *Transportation Research Part F: Traffic Psychology and Behaviour*, *104*, 474–487. https://doi.org/10.1016/j.trf.2024.06.014

Wang, A., Wang, J., Shi, W., & He, D. (2024). Cognitive Workload Estimation in Conditionally Automated Vehicles Using Transformer Networks Based on Physiological Signals. *Transportation Research Record*. https://doi.org/10.1177/03611981241250023

Wang, J., Tu, R., Wang, A., & He, D. (2024). Trust in Range Estimation System in Battery Electric Vehicles–A Mixed Approach. *IEEE Transactions on Human-Machine Systems*, *54*(3), 250–259. https://doi.org/10.1109/THMS.2024.3381116

Wang, J., Yang, X., Wang, Z., Wei, X., Wang, A., He, D., & Wu, K. (2024). *Efficient Mixture-of-Expert for Video-based Driver State and Physiological Multi-task Estimation in Conditional Autonomous Driving* (arXiv:2410.21086). arXiv. https://doi.org/10.48550/arXiv.2410.21086

Widyanti, A., Hanna, Muslim, K., & Sutalaksana, I. Z. (2017). The sensitivity of Galvanic Skin Response for assessing mental workload in Indonesia. *Work*, *56*(1), 111–117. https://doi.org/10.3233/WOR-162479

Wu, G., Wang, A., Huang, C., & He, D. (2023). Phone call while riding among delivery workers: The Application of the Theory of Planned Behavior. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. https://doi.org/10.1177/21695067231193658

Wu, H., Wu, C., Lyu, N., & Li, J. (2022). Does a faster takeover necessarily mean it is better? A study on the influence of urgency and takeover-request lead time on takeover performance and safety. *Accident Analysis & Prevention*, *171*, 106647. https://doi.org/10.1016/j.aap.2022.106647

Yang, H., Zhou, Y., Wu, J., Liu, H., Yang, L., & Lv, C. (2025). Human-Guided Continual Learning for Personalized Decision-Making of Autonomous Driving. *IEEE Transactions on Intelligent Transportation Systems*, 1–0. IEEE Transactions on Intelligent Transportation Systems. https://doi.org/10.1109/TITS.2024.3524609

Yao, H., An, S., Zhou, H., & Itoh, M. (2021). Driver takeover performance in conditionally automated driving: Sudden system failure situation versus ODD exit situation. *SICE Journal of Control, Measurement, and System Integration*, *14*(1), 89–96. https://doi.org/10.1080/18824889.2021.1913861

Yi, B., Cao, H., Song, X., Wang, J., Zhao, S., Guo, W., & Cao, D. (2023). How Can the Trust-Change Direction be Measured and Identified During Takeover Transitions in Conditionally Automated Driving? Using Physiological Responses and Takeover-Related Factors. *Human Factors*. https://doi.org/10.1177/00187208221143855

Yoon, S. H., & Ji, Y. G. (2019). Non-driving-related tasks, workload, and takeover performance in highly automated driving contexts. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 620–631. Scopus. https://doi.org/10.1016/j.trf.2018.11.015

Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis & Prevention*, *92*, 230–239. https://doi.org/10.1016/j.aap.2016.04.002

Zhu, J., Ma, Y., Zhang, Y., Zhang, Y., & Lv, C. (2023). Takeover quality prediction based on driver physiological state of different cognitive tasks in conditionally automated driving. *Advanced Engineering Informatics*, *57*, 102100. https://doi.org/10.1016/j.aei.2023.102100