Effect of Riding Experience and HMI on Users' Trust and Riding Comfort in Fully Driverless Autonomous Vehicles: An On-Road Study

Weiyin Xie[a], Zhenyu Wang[b], Dengbo He[a,b,*]

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

[b] Thrust of Intelligent Transportation, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China

[*]Corresponding author. Email: dengbohe@hkust-gz.edu.cn

ABSTRACT

The wide adoption of autonomous vehicles (AVs) or robot taxis relies on technological advancements and public acceptance, which can be influenced by users' trust in AVs and comfort during rides. Among the influential factors of riding comfort, motion sickness (MS) has attracted lots of attention in previous research, and both trust and MS have been found to be associated with human-machine interface (HMI) designs in AVs. However, previous research on trust and MS in AVs predominantly utilized driving simulations or "Wizard of Oz" methods, which failed to introduce risk and realistic vehicle motions, potentially introducing bias to conclusions. For the first time, our study investigated the impact of displaying the dynamic path trajectories of AVs on passengers' perceptions of system transparency, trust, and MS in a commercially running AV. The results from 16 participants and 32 rides revealed limited effects of the dynamic path trajectory on trust, and a discernible but statistically non-significant trend in motion sickness alleviation. Further, we found that the initial riding experience was more important in trust enhancement than subsequent rides. These results provide insights into future HMI design in robot taxis and suggest directions for future research in trust enhancement and MS alleviation in AVs.

Keywords: motion sickness; trust; on-road study; autonomous vehicle

1. Introduction

With the rapid advancement of driving automation technologies, Level 4 autonomous vehicles (AVs) have become commercially operational in several cities (Nina, 2022; SAE, 2021). The widespread adoption of AVs is anticipated to generate substantial economic and societal benefits by enabling passengers to repurpose commuting time for work or entertainment activities. However, two critical challenges hinder the mass adoption of AVs: users' relatively low trust in autonomous systems (Hegner et al., 2019) and the high incidence of motion sickness (MS) during engagement in non-driving-related tasks (NDRTs) while riding in AVs (Isu et al., 2014).

Trust is a key determinant of users' acceptance of AV technology (Adnan et al., 2018) and is influenced by various factors, including driving style (Ekman et al., 2019), individual characteristics (Mosaferchi et al., 2023), and the design of human-machine interfaces (HMIs) (Qi et al., 2024). Among these, providing HMIs to users has been identified as a controllable and low-cost solution. Thus, previous research and operators of AV fleets have proposed different human-machine interfaces (HMIs) to improve users' trust in AVs. For example, Oliveira et al. (2020) found that providing information regarding an AV's intended trajectory and the perception of hazards can boost passengers' trust in AVs. Another study investigating the impact of HMI design on trust had similar findings, indicating that presenting driving-related information through HMI can enhance passengers' trust in AVs (Hartwich et al., 2021). However, these findings were mostly based on studies conducted in simulated environments or "Wizard of Oz" conditions, where actual driving risks and realistic vehicle dynamics are absent. Consequently, the external validity of these findings remains uncertain, underscoring the need for validation under real-world traffic conditions.

At the same time, MS, primarily caused by sensory conflict (Reason, 1975), presents another significant barrier to AV adoption. As AVs liberate users from monotonous driving tasks and transition them into the role of passengers, engagement in non-driving-related tasks (NDRTs), such as watching videos, becomes common. However, passengers, unlike drivers, are more susceptible to MS, likely due to their reduced ability to anticipate future vehicle movements (Rolnick & Lubow, 1991). Furthermore, the introduction of in-vehicle displays may exacerbate MS symptoms, especially when passengers are engaged in visually demanding tasks such as reading (Morimoto et al., 2008) or video watching (Kato & Kitazaki, 2006). Prior research suggests that providing anticipatory information about future vehicle maneuvers, such as upcoming turns, accelerations, or decelerations, can help alleviate MS symptoms (Karjanto et al., 2018; Li & Chen, 2022; Maculewicz et al., 2021). However, most existing studies employed static, discrete cues delivered several seconds before the actual vehicle motion. For example, in previous field experiments, motion planning information was pre-defined and consistently matched the upcoming maneuver, with signals such as LED stripes indicating a lane change 1–3 seconds in advance. While effective in controlled environments, such setups are unrealistic in real-world driving, where an AV's path planning must be continuously updated in response to dynamic traffic conditions. Consequently, the potential of real-time dynamic trajectory information to mitigate MS remains underexplored.

Thus, displaying upcoming trajectories of AVs on the in-vehicle HMI seems to benefit AV occupants from both trust and MS perspectives, which has already been adopted by some AV fleet operators, such as robotaxis operated by Pony.ai and WeRide (Nina, 2022; Nolan, 2025). Despite the potential benefit, the influence of real-time dynamic trajectory plan information on passengers' trust in AVs and their experience of MS has not been empirically validated in commercially operating AVs. To address this gap, we investigated the influence of real-time dynamic trajectory plan information on passengers' trust and MS development in a commercially operating fully driverless taxi service. Given that the initial riding experience plays an important role in shaping one's trust in AVs (Wang et al., 2024), we focused on participants who had no prior experience in taking a ride in AVs.

Accordingly, three research questions (RQs) were proposed. RQ1: Can displaying the dynamic trajectory plan in real time increase occupants' trust in AVs? RQ2: Can displaying dynamic trajectory plan information help alleviate MS when occupants are engaged in visual NDRTs in AVs? RQ3: Does riding experience increase users' trust in AVs? Three corresponding hypotheses were formulated regarding the RQs, i.e., H1: Providing passengers with a real-time dynamic trajectory plan can increase their trust in AVs, as such information may increase system transparency, which is positively associated with users' trust in automation (Yang et al., 2017). H2: A real-time dynamic trajectory plan can help alleviate MS among passengers in AVs, as it can support passengers' expectations of AV's future motions. H3: The riding experience is positively associated with users' trust in AVs.

2. Methodology

2.1 Participants

A total of 16 participants (eight males and eight females) with an average age of 22.7 years (range: 18-32, standard deviation (SD) = 4.1) were recruited through online forums and campus posters. Participants were required to self-report with no prior riding experience in fully driverless AVs or AVs with safety guards. Additionally, they needed to have self-reported MS susceptibility scores above the 50th percentile (i.e., scoring at least 11.25 out of 54 points on the Motion Sickness Susceptibility Questionnaire Short-form [MSSQ-short] (Golding, 2006) to ensure they are prone to MS in vehicles. The Chinese version of the MSSQ-short was utilized, as the experiment took place in mainland China (Lin & Guo, 2022). The participants' mean MSSQ score was 19.6 (range: 12-29, SD = 5.8), and all reported good health and normal vestibular functions.

2.2 Experiment Design

Participants engaged in a closed-loop AV ride beginning at the Guangzhou campus of the Hong Kong University of Science and Technology (HKUSTGZ), passing through Huangge Automobile City, and concluding back at the HKUSTGZ (see Figure 1). The AV had to decelerate and accelerate to follow the traffic or respond to traffic lights at intersections, and change lanes if necessary to overtake slow-moving traffic. Thus, this route led to a typical urban driving trajectory involving both longitudinal and lateral accelerations. To mitigate the influence of traffic flow, the rides occurred exclusively on weekdays between 9:00 am and 4:00 pm, avoiding peak hours. In all trials, participants took the same type of 7-seat MPV provided by the same robotaxi service provider (see Figure 2a). The intended trajectory and perceived traffic were originally supplied by the service provider and shown on the headrest screen. In addition, the screen also showed the perceived environment in real time, including all detected road agents (i.e., motor vehicles, non-motor vehicles, and pedestrians) and key road or traffic elements (e.g., traffic lights, lane markings, curbs, and temporary traffic cones). However, information regarding the ego-vehicle velocity was not displayed during the experiments.

In the experiment, a customized screen cover was utilized to conceal this information in half of the rides (see Figure 2c), while it was removed in the other half of the rides, allowing passengers to view the dynamic trajectory plan of the AV (see Figure 2d). Vehicle dynamics were recorded using an Xsens Inertial Measurement Unit (IMU) mounted at a fixed position on the left armrest of the right-side seat in the second row (Figure 2b). The IMU sampled linear acceleration at 60 Hz with an accuracy of 0.01 G. According to the right-hand coordinate system specified in ISO8855 (2011), the z-, x-, and y-axes corresponded to the upward (vertical), forward (longitudinal), and rightward (lateral) directions of the vehicle, respectively.

The order of the experiment conditions, that is, the trajectory HMI availability (i.e., with and without), was adopted as the within-subject factor and was fully counterbalanced to reduce the learning effect (Figure 3). To minimize the aftereffects of MS, we ensured at least a 3-day gap between the two trials of the same participant.

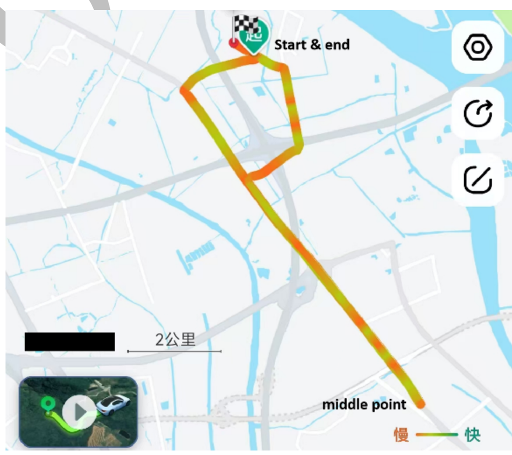
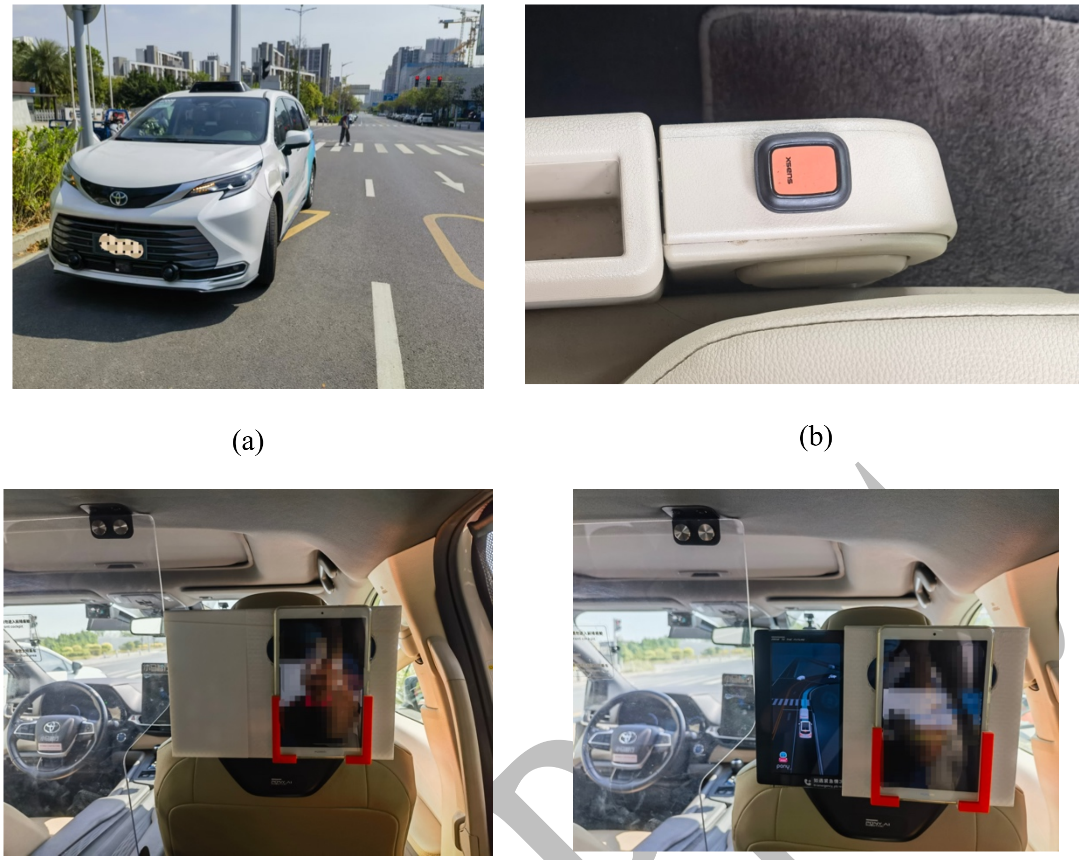


Figure 1. The 15-km route for each ride.



(a) (b)

(c) (d)

Figure 2. Apparatus: (a) The robotaxi used in this experiment; (b) Xsens IMU for recording vehicle dynamics; (c) Experimental condition without trajectory plan HMI; (d) Experimental condition with trajectory plan HMI.

At the same time, throughout each ride, the in-vehicle display showed documentary videos regarding food and history as the sole task permissible for participants. The videos were presented in Chinese, the participants' native language, to reduce comprehension difficulties. Participants were informed that they needed to answer questions related to the videos after rides, and the correctness of the answers would influence the experimental compensation, aiming to encourage them to focus on this visual NDRT. During the ride, an experimenter accompanied the participant, sitting in the left seat of the second row, and occasionally reminded the participant to focus on the video task if they became distracted by the external environment. Participants consistently occupied the right seat of the second row. While the video content differed between rides for each participant to ensure attentiveness, the sequence of the videos remained consistent for all participants to neutralize potential video content effects on MS (see Figure 3).

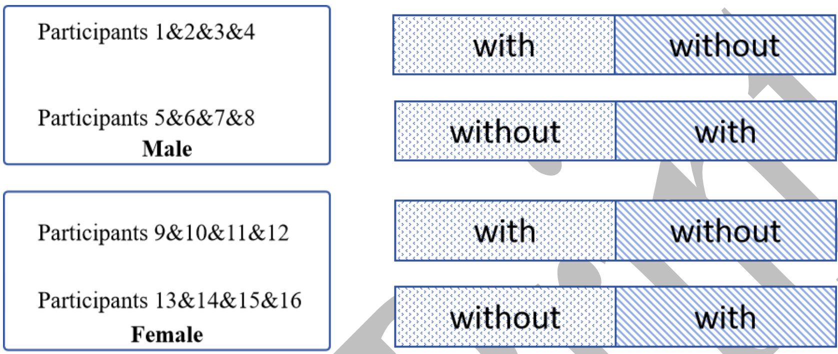


Figure 3. The order of the experimental conditions (with or without trajectory HMI) and the video content. The blocks with the same background pattern indicate identical video content.

2.5 Experiment Procedures

Following Suwa et al. (2022), participants were instructed to prioritize self-care, ensure adequate sleep, abstain from alcohol at least one day before the experiment, and finish meals at least an hour before the experiment to maintain optimal health conditions. Upon arrival on the experiment day, participants signed a consent form regarding the objectives, procedures, risks, and benefits of the experiment. Before the ride, participants completed a pre-trial 7-point trust questionnaire (Manchon et al. (2022); see Table A.1 in the Appendix) and an MS assessment questionnaire (MSAQ) (Gianaros et al., 2001). They were also provided with an explanation of the Chinese version of the Misery Scale (MISC) (Lin & Guo, 2022), which measures the level of MS (Bos et al. (2005); see Table 1). During the ride, participants were required to watch video clips on an in-vehicle display. At the same time, a timer emitted an auditory alert at one-minute intervals (90 dB loudness, 1000 Hz frequency, 1-second duration) to prompt participants to orally report their MS levels using the MISC. To assist them in recalling the scale meanings, the MISC was printed and affixed to the frame of the display screen. After each ride, they completed the same trust questionnaire and the MSAQ again. Additionally, they completed a 7-point system transparency questionnaire (Choi & Ji, 2015) to assess their perceived system transparency of the AV (see Figure 4 and Table 2), both before and after the ride. The MSAQ has four constructs that categorize MS symptoms into four distinct dimensions: gastrointestinal, central, peripheral, and sopite-related. This experiment received approval from the Human and Artefacts Research Ethics Committee of the Hong Kong University of Science of Technology (HREP-2023-0246).

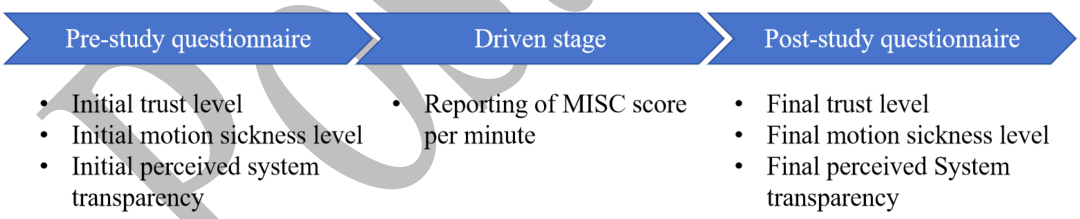


Figure 4. The experimental procedures.

Table 1. MISC for MS evaluation during the rides

Symptoms		MISC
No problems		0
Some discomfort, but no specific symptoms		1
Dizziness, cold/warm, headache, stomach/throat awareness, sweating, blurred vision, yawning, burping, tiredness, salivation, but no nausea	Vague	2
	Little	3
	Rather	4
	Severe	5
Nausea	Little	6
	Rather	7
	Severe	8
	Retching	9
Vomiting		10

Table 2. Perceived system transparency questionnaire

Q1	I believe that the AV acts consistently and its behavior can be forecast.
Q2	I believe that I can form a mental model and predict the future behavior of the AV.

2.6 Statistical Analysis

In this study, a two-step data analysis was conducted in *R* (4.4.1).

Step 1: Quantification of motion stimuli and NDRT engagement

First, motion sickness dose values (MSDV) were used to compare motion stimuli in both experimental conditions (i.e., with and without the dynamic trajectory plan), adhering to

ISO2631-1 (1997) guidelines. The comprehensive $MSDV_{total}$, as defined by Yunus et al. (2021) and expressed in Equation (1), quantified the motion stimuli experienced during a ride.

$$\boldsymbol{MSDV_{total}} = \sqrt{\boldsymbol{MSDV_x}^2 + \boldsymbol{MSDV_y}^2 + \boldsymbol{MSDV_z}^2} \qquad (1)$$

$MSDV_x$, $MSDV_y$, and $MSDV_z$ calculations, as detailed in Equations (2) and (3), involved the weighted integration of acceleration along different directions, where $\boldsymbol{W_i}$ indicated the acceleration weight. The frequency weighting specified in ISO2631-1 (1997) was applied to compute $MSDV_z$ in the vertical direction, and frequency weightings from Donohew and Griffin (2004) were used to calculate $MSDV_x$ and $MSDV_y$ in the lateral and longitudinal directions, respectively.

$$\boldsymbol{a_{i,w}(t)} = \boldsymbol{a_i(t)} \times \boldsymbol{W_i} \qquad \boldsymbol{i = x, y, z} \qquad (2)$$

$$\boldsymbol{MSDV_i} = \left(\int_0^T (\boldsymbol{a_{i,w}(t)})^2\right)^{1/2} \qquad (3)$$

To ensure that differences in participants’ engagement with the video-watching task did not confound the development of motion sickness, we conducted an additional paired t-test on participants' video task performance across conditions.

Step 2: Analysis of perceived transparency and trust

To assess subjective responses, linear mixed-effects models were fitted using the lmer function in *R* (Marc, 2015). Two separate models were specified:

- Model 1 for perceived system transparency
- Model 2 for trust in the system

As shown in Table 3, the fixed effects in both models included HMI availability (with vs. without), measurement timing (pre- vs. post-ride), ride order (first vs. second ride), and their two-way interactions. Both models also incorporated random intercepts and slopes to account for individual differences.

Table 3. The structure of the models for perceived transparency and trust

Model	Dependent Variable	Fixed Effects	Random Effects
Model 1	Perceived Transparency	- HMI availability - Timing (pre/post) - Ride number - HMI availability× Ride number - Timing × Ride number - HMI availability× Timing	- Intercept for participant - Slopes for (all nested in Participant): • Ride number • HMI availability • Timing
Model 2	Trust Score	- HMI availability - Timing (pre/post) - Ride number - HMI availability× Ride number - Timing × Ride number - HMI availability× Timing	- Intercept for participant - Slopes for (all nested in Participant): • Ride number • HMI availability • Timing

Step 3: Analysis of motion sickness outcomes

As shown in Table 4, two sets of analyses were conducted to evaluate MS symptoms. First, a linear mixed-effects model (Model 3) was used to examine the effects of HMI availability, timing, and their interaction on the MSAQ scores. Due to non-normality, a square root transformation was applied to the MSAQ data to meet ANOVA assumptions.

For significant independent variables ($p < .05$), we used the "emmeans" package in R (Lenth et al., 2024) for post hoc comparisons, with the “Tukey” method adopted for p-value adjustment. Furthermore, to gain insight into how each construct of the MSAQ was influenced by the experimental conditions (i.e., with and without the dynamic trajectory plan information), Wilcoxon Signed-Rank Tests (WSRT) were conducted for the four constructs of MSAQ, given the non-normal distribution of the data.

Table 4. The structure of the models for motion sickness measures

Model	Dependent Variable	Fixed Effects	Random Effects
Model 3	sqrt (MSAQ)	- HMI availability - Timing (pre/post) - HMI availability × Timing	- Intercept for participant - Slopes for (all nested in Participant): • HMI • Timing
Mode 4	MISC score (per minute)	- HMI availability - Time of measurement (minute by minute) - HMI availability×Time of measurement	- Intercept for participant

Second, a linear mixed-effects model (Model 4) was built to analyze the minute-by-minute fluctuation of MISC scores during each ride. Time of measurement (in minutes), HMI availability, and their interaction were used as the fixed effects, with a participant-level random intercepts to account for repeated measures on each participant.

3. Results

3.1 Evaluation of Experimental Control

The average ride duration across all trials was 29.7 minutes (SD = 4.1 min; range: 24–42 min; see Figure 9). The Wilcoxon rank-sum test revealed no significant difference in trial duration between the two conditions (p > 0.05). To ensure fair comparisons, we first compared the effectiveness of the experiment control between the HMI conditions.

First, to ensure that the differences in MS were not due to the road profiles across drives, we conducted a t-test to compare motion stimuli (i.e., MSDV) between the two conditions (see Figure 5). The analysis revealed no significant difference in MSDV between the two conditions ($p > .05$).

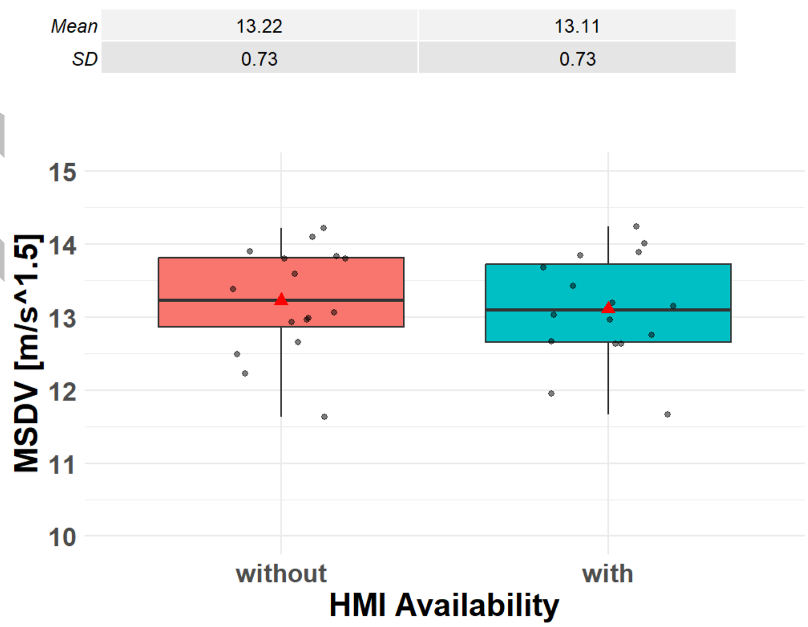


Figure 5. MSDV in both experimental conditions (with versus without a dynamic trajectory HMI)

Second, the performance in the video-watching task under each condition was also compared. Specifically, the accuracies of participants' responses to video-related questionnaires were both 75%, regardless of HMI availability, with an SD of 24.2% and 27.4% with and without the HMI, respectively. The paired t-test revealed no significant differences across experimental conditions ($p > .05$), suggesting no significant variations in the level of video-watching task engagement between the experimental conditions.

3.1 System Transparency and Trust

Table 5 summarizes the results of the repeated measures ANOVA, with the transparency score, trust score, and overall MSAQ score as the dependent variables (DVs) and the results of the repeated measures ANOVA, with the trust score as the DV. As shown in Figure 6, the post hoc comparison of the timing of measurements indicates that perceived system transparency increased by 1.46 post-ride compared to pre-ride, with a 95% confidence interval (95% CI) of [1.17, 1.75] and $t(45) = 10.09$.

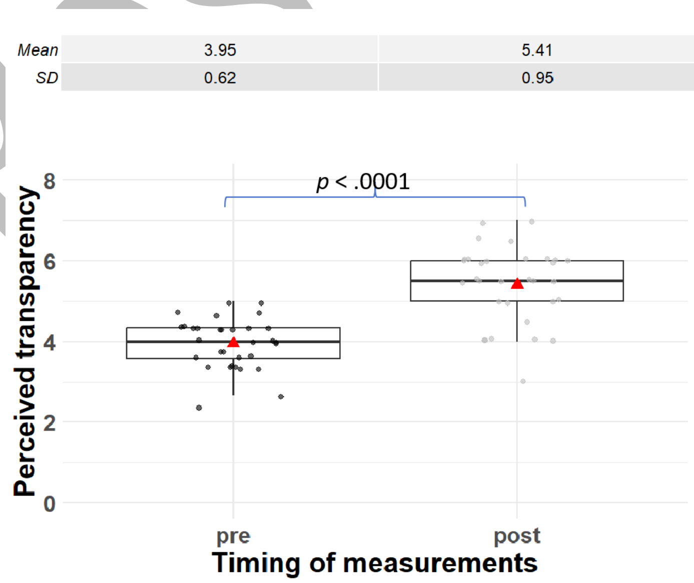


Figure 6. Effect of the timing of measurements on the perceived transparency

Table 5. Results of the statistical models.

	HMI availability	Timing	Number of rides	HMI availability × timing	Timing × Number of rides	HMI availability × Number of rides	Time of measurement	HMI availability × Time of measurement
Transparency	F(1,28)=0.27 *p* = .6	F(1,15)=87.92 *p* < .0001*	F(1,28)=0.44 *p* = .5	F(1,28)=1.96 *p* = .2	F(1,28)=0.44 *p* = .5	F(1,14)=0.82 *p* = .4	-	-
Trust	F(1,14)=1.12 *p* = .3	F(1,15)=17.48 *p* = .0008*	F(1,14)=6.85 *p* = .02*	F(1,14)=0.97 *p* = .3	F(1,14)= .08 *p* = .04*	F(1,14)=1.59 *p* = .2	-	-
Overall MSAQ	F(1,45)=0.24 *p* = .6	F(1,45)=26.87 *p* < .0001*	-	F(1,45)=0.18 *p* = .7	-	-	-	-
MISC score (per minute)	F(1,931.22)=0.0 *p* = .9	-	-	-	-	-	F(1,931.46)=194.5 *p* < .0001*	F(1,931.59)=0.7 *p* = .4

*Note: * indicates significant results (p < .05)*, and "-" means not applicable.

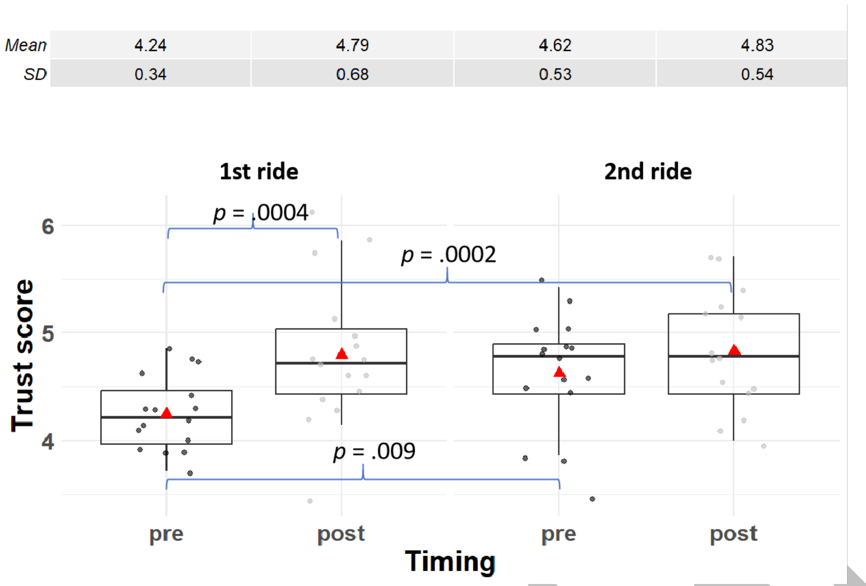


Figure 7 Interaction effects of measurement timing and number of rides on the trust score

Figure 7 illustrates the post hoc comparison of the interaction effects of the number of rides and measurement timing on the trust score. The results indicate that, for the first ride, the post-ride trust score increased by 0.55 units compared to the pre-ride trust score, with a 95% confidence interval (95% CI) of [0.23, 0.88] and a t-value of $t(28.6) = 4.65$. At the same time, the pre-ride trust score was higher in the second ride as compared to the first ride, with a difference (Δ) of 0.38 units, 95% CI: [0.08, 0.69], $t(28) = 3.44$ and the post-ride trust score of the second ride was higher than the pre-ride trust score of the first ride, $\Delta = 0.59$ units, 95% CI of [0.26, 0.92], $t(28.8) = 4.87$.

3.2. Motion sickness

3.2.1. MSAQ score

The results of the linear mixed-effects model indicated a significant main effect of timing, whereas the interaction between timing and the number of rides was non-significant ($p > .05$, see Table 6). Thus, a post hoc analysis was performed for the significant timing effect. As shown in Figure 8, we observed a significant increase in the square root of the MSAQ score from pre-ride to post-ride, Δ = 1.47, 95% CI of [0.90, 2.05], $t(45) = 5.16$.

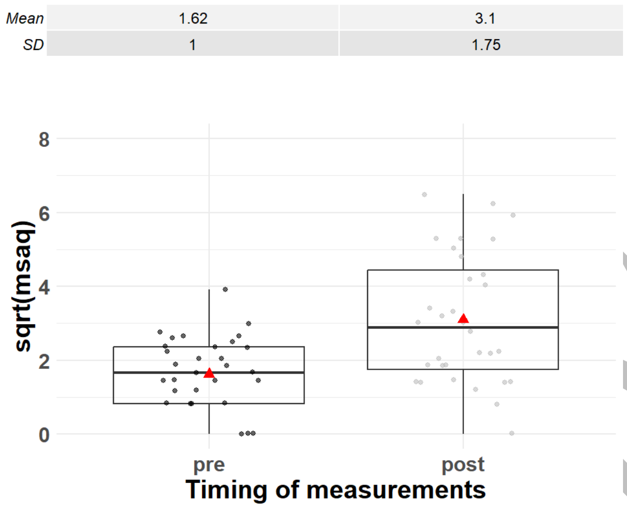


Figure 8 Effect of timing of measurements on the root of the square of the MSAQ score

The Wilcoxon Signed-Rank Test was conducted to further compare the MSAQ scores for each construct (see Table 6). The results showed that without the HMI, significant increases were observed in gastrointestinal, central, and sopite-related scores, but not in peripheral-related scores; whereas with the HMI, only the central and sopite-related scores showed significant increases.

Table 6. Wilcoxon Signed-Rank Test (WSRT) results for the pre-ride and post-ride MSAQ

Condition	MSAQ		Median (IQR)	WSRT
Without a dynamic trajectory plan HMI	G	Pre	0.00 (0.00-0.00)	**z = -2.50, r = 0.70, *p* = .01***
		Post	4.17 (0.00-28.47)	
	C	Pre	0.00 (0.00-2.22)	**z = -2.86, r = 0.72, *p* = .004***
		Post	11.11 (1.11-16.11)	

	P	Pre	0.00 (0.00-12.04)	z = -0.28, r = 0.07, *p* = .8
		Post	3.70 (3.70-8.33)	
	S	Pre	4.17 (0.00-9.72)	**z = -2.52, r = 0.63, *p* = .01***
		Post	13.89 (5.56-28.47)	
	O	Pre	3.13 (1.22-5.73)	**z = -2.80, r = 0.70, *p* = .005***
		Post	9.72 (4.51-23.61)	
With a dynamic trajectory plan HMI	G	Pre	0.00 (0.00-3.47)	z = -1.67, r = 0.42, *p* = .10
		Post	2.78 (0.00-18.06)	
	C	Pre	0.00 (0.00-2.22)	**z = -2.85, r = 0.71, *p* = .004***
		Post	4.44 (2.22-11.67)	
	P	Pre	1.85 (0.00-11.11)	z = -0.27, r = 0.07, *p* = .8
		Post	3.70 (0.00-15.74)	
	S	Pre	4.17 (0.00-8.33)	**z = -2.49, r = 0.62, *p* = .01***
		Post	9.72 (2.78-20.14)	
	O	Pre	2.43 (0.69-4.51)	**z = -2.30, r = 0.58, *p* = .02***
		Post	5.56 (2.08-16.32)	

Note: In this and subsequent tables, IQR denotes the inter-quartile range. The abbreviations G, C, P, and S correspond to the gastrointestinal, central, peripheral, and sopite-related MSAQ constructs, respectively, while O represents the overall MSAQ score.

Moreover, another Wilcoxon Signed-Rank Test was performed to compare the difference in the change of MSAQ scores (i.e., the difference between scores measured before and after each ride) between both conditions (see Table 7). The results showed no significant differences in changes in MSAQ constructs between conditions with and without trajectory plan information. Nevertheless, a noticeable trend emerged: the construct scores increased to a lesser extent in the treatment conditions compared to the control conditions.

Table 7. Wilcoxon Signed-Rank Test results for the difference in the change between pre- and post-ride MSAQ scores across both conditions (with versus without HMI availability)

MSAQ	Condition	Median (IQR)	WSRT
G	Without	4.17 (0.00-22.22)	z = -0.90, r = 0.23, p = .4
	With	0.00 (0.00-11.11)	
C	Without	10.00 (0.00-16.11)	z = -1.13, r = 0.28, p = .3
	With	2.22 (2.22-9.44)	
P	Without	0.00 (-8.33-3.70)	z = -0.91, r = 0.23, p = .4
	With	0.00 (-7.41-3.70)	
S	Without	8.33 (2.08-15.28)	z = -1.27, r = 0.32, p = .2
	With	5.56 (2.78-12.50)	
O	Without	6.94 (2.60-17.19)	z = -0.77, r = 0.19, *p* = .4
	With	2.08 (0.35-10.24)	

3.2.2. MISC score

Table 5 summarizes the results of the linear mixed-effects model, with the MISC score measured per minute as the dependent variable. Figure 9 illustrates the fluctuation in MISC scores for each participant throughout the ride under the two experimental conditions. Results from the linear mixed-effects model revealed no significant effect of the HMI condition on MISC scores (see Table 5).

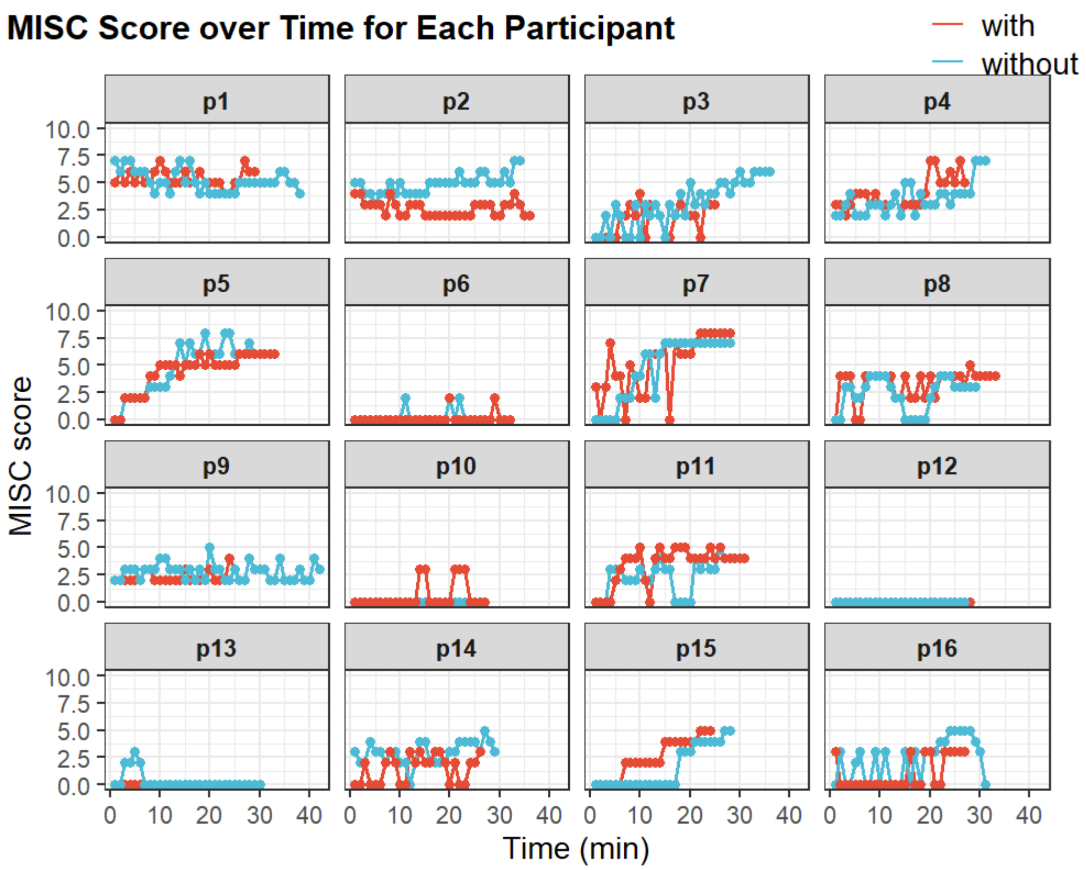


Figure 9. MISC score of all 16 participants (p1-p16) during the ride in both HMI conditions.

4. Discussion

Despite extensive research on trust and MS through driving simulators or Wizard of Oz methodologies, our study explores the impact of the dynamic trajectory plan information on trust enhancement and MS reduction in commercially running AVs.

First, in contrast to previous findings from the Wizard of Oz approach (Oliveira et al., 2020), we found that the dynamic trajectory information failed to contribute to users' perceived system transparency or trust enhancement. This suggests that, contrary to our hypothesis (H1), the effectiveness of this type of HMI design may be limited. These results may be explained from two perspectives. First, our participants were predominantly college students who are overall young. This population was found to have a better understanding of

AV technology and potentially higher initial perceptions of system transparency as compared to the older and less educated population (Thomas et al., 2020). Thus, it is challenging to further boost their perceived transparency of the system. Second, throughout the whole drive, the vehicle operated well, and all participants experienced smooth rides, thereby diminishing the observable impact of HMI design on the perceived transparency and trust of the AVs.

Further, we failed to observe the influence of dynamic trajectory information on MS development in terms of the overall MASQ score and the MISC score (H2). However, upon further examination of the MSAQ construct, we found that the overall increase in MSAQ scores in rides across all constructs was smaller when trajectory information was displayed compared to when it was not. Further, with the HMI, the increase of several constructs became nonsignificant after rides. These findings suggest that the dynamic trajectory plan information has the potential to alleviate MS, but the magnitude of this alleviation is relatively small. It is also possible that, unlike human drivers, the AVs adopted a relatively conservative driving style (Wenger, 2024) and rarely conducted abrupt maneuvers, especially in the lateral direction. While displaying trajectory information can assist passengers in anticipating lateral acceleration during lane changes or turns, it does not convey information on longitudinal acceleration experienced during lane following. Conveying longitudinal acceleration information has been found to be able to alleviate MS in previous studies (Kehl et al., 2024). This suggests that for future HMI designs aimed at mitigating MS in AVs, acceleration data should also be integrated into the trajectory visualization. For instance, varying colors can be used to represent different levels of acceleration, similar to how levels of sound loudness can alleviate MS (Xie et al., 2025).

Importantly, the limited effects observed in this study may be closely tied to the specific HMI design used in this study. The current interface provides only action-based information (e.g., where the vehicle is going) without explaining the rationale behind these actions. Prior research suggests that explaining the reasons behind system behaviors is critical for enhancing system transparency and fostering trust (Luo et al., 2022). Incorporating such reasoning into future HMI designs may enhance users' trust in AVs. Further, the trajectory information was displayed only at the periphery of the screen, while participants were visually engaged with an NDRT. This may have led them to overlook or pay limited attention to the HMI. At the same time, previous studies have shown that visual anticipatory cues are generally less effective at mitigating MS than non-visual cues such as auditory or tactile feedback (Xie et al., 2023). This may be because the visual channel is often occupied with other stimuli, whereas non-visual channels can more effectively capture attention via preemption effects (Smith et al., 2009; Wickens et al., 2005). Thus, future designs could consider combining visual trajectory displays with auditory or vibrotactile cues to improve both MS mitigation and trust formation.

Finally, our findings support the effectiveness of the first ride in trust enhancement (H3), and the increase becomes insignificant for the second ride. This suggests that the first riding experience plays a more crucial role in enhancing trust than subsequent rides. After the initial increase, trust levels stabilized. The benefit of the first ride experience on the users' trust in

AV has been observed in a previous study (Wang et al., 2024). Our findings further confirm that this increase in trust may last and can be further enhanced by following rides. This increase in trust may have been enhanced by the increase in the perceived system transparency during the rides, as we also observed a significant increase in perceived system transparency from pre-ride to post-ride. This finding holds important implications for AV service providers: they should prioritize creating a positive and comfortable first ride for passengers to increase the adhesiveness of users.

5. Limitations

This research has several limitations. Firstly, the statistical power of our experimental design is optimized for detecting large effect sizes ($f^2 \geq 0.4$, alpha = 0.05, power =0.8), which means that smaller effect sizes may not be captured within our framework. Second, all participants in this study were from the younger generation, who typically have a better understanding of new technology, resulting in higher initial perceptions of system transparency and trust. Future research should aim to include a larger sample size and a more diverse demographic representation. Finally, the dynamic trajectory plan information in this study was presented using animation, along with other detected road elements (e.g., surrounding vehicles and traffic signs). Though it is a common practice to provide both the planned trajectory and the perceived environment simultaneously in robot taxis, future studies could explore other rendering techniques, such as augmented reality (AR), and differentiate the effect of the perceived road element and the trajectory on trust formation and MS development in AVs.

6. Conclusion and Applications

In this study, we investigated the effects of dynamic trajectory plan information and riding experiences on trust enhancement and MS alleviation in a fully driverless autonomous vehicle. Our findings can be summarized as follows:

- We found no significant effects of trajectory plan information on perceived system transparency and trust enhancement. As such, the visualization of the dynamic trajectory plan may be considered optional in commercially running AVs, from users' trust and perceived transparency perspectives of view.
- A Wilcoxon Signed-Rank Test of variations in MSAQ constructs suggests that trajectory information may alleviate MS, although simply displaying the planned trajectory without conveying acceleration yielded insignificant results in ANOVA analysis.
- The initial riding experience plays a critical role in enhancing trust than subsequent rides, and trust does not appear to increase with one additional ride. AV operators may put more emphasis on attracting first-time users.

Author Contribution

Weiyin Xie: Conceptualization, Data curation, Formal analysis, Writing – original draft, Methodology, Investigation; Zhenyu Wang: Data curation, Writing – review & editing; Dengbo He: Funding acquisition, Project administration, Supervision, Writing – review & editing

Declaration of Competing Interest

The authors acknowledge that they have no conflicts of interest in this work.

Acknowledgments

This research is supported by the Guangzhou-HKUST(GZ) Joint Funding Scheme (No. 2023A03J0150), and partially by the Guangzhou Municipal Science and Technology Project (No. 2023A03J0011), and the Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No. 2023B1212010007).

REFERENCES

Adnan, N., Nordin, S. M., bin Bahruddin, M. A., & Ali, M. (2018). How trust can drive forward the user acceptance to the technology? In-vehicle technology for autonomous vehicle. *Transportation Research Part A: Policy and Practice*, *118*, 819-836.

Bos, J. E., MacKinnon, S. N., & Patterson, A. (2005). Motion sickness symptoms in a ship motion simulator: effects of inside, outside, and no view. *Aviation, Space, and Environmental Medicine*, *76*(12), 1111-1118.

Choi, J. K., & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction*, *31*(10), 692-702.

Donohew, B. E., & Griffin, M. J. (2004). Motion sickness: effect of the frequency of lateral oscillation. *Aviation, Space, and Environmental Medicine*, *75*(8), 649-656.

Ekman, F., Johansson, M., Bligård, L.-O., Karlsson, M., & Strömberg, H. (2019). Exploring automated vehicle driving styles as a source of trust information. *Transportation Research Part F: Traffic Psychology and Behaviour*, *65*, 268-279.

Gianaros, P. J., Muth, E. R., Mordkoff, J. T., Levine, M. E., & Stern, R. M. (2001). A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, Space, and Environmental Medicine*, *72*(2), 115.

Golding, J. F. (2006). Predicting individual differences in motion sickness susceptibility by questionnaire. *Personality and Individual Differences*, *41*(2), 237-248.

Hartwich, F., Hollander, C., Johannmeyer, D., & Krems, J. F. (2021). Improving passenger experience and Trust in Automated Vehicles through user-adaptive HMIs:“The more the better” does not apply to everyone. *Frontiers in Human Dynamics*, *3*, 669030.

Hegner, S. M., Beldad, A. D., & Brunswick, G. J. (2019). In automatic we trust: investigating the impact of trust, control, personality characteristics, and extrinsic and intrinsic motivations on the acceptance of autonomous vehicles. *International Journal of Human–Computer Interaction*, *35*(19), 1769-1780.

ISO2631-1. (1997). Mechanical vibration and shock: Evaluation of human exposure to whole-body vibration—Part 1: General requirements (ISO Standard No. 2631-1). In: International Organization for Standardization.

ISO8855. (2011). Road vehicles—Vehicle dynamics and road-holding ability—Vocabulary (ISO Standard No. 8855:2011). In: International Organization for Standardization.

Isu, N., Hasegawa, T., Takeuchi, I., & Morimoto, A. (2014). Quantitative analysis of time-course development of motion sickness caused by in-vehicle video watching. *Displays*, *35*(2), 90-97.

Karjanto, J., Yusof, N. M., Wang, C., Terken, J., Delbressine, F., & Rauterberg, M. (2018). The effect of peripheral visual feedforward system in enhancing situation awareness and mitigating motion sickness in fully automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *58*, 678-692.

Kato, K., & Kitazaki, S. (2006). *A study for understanding carsickness based on the sensory conflict theory* (SAE Technical Paper No. 2006-01-0096).

Kehl, L., Brietzke, A., Xuan, R. P., & Hecht, H. (2024). Anticipatory vibrotactile cues reduce motion sickness in car passengers during stop-and-go driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *107*, 791-802.

Lenth, R., Bolker, B., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Piaskowski, J., Riebl, H., & Singmann, H. (2024). *Estimated marginal means, aka least-squares means*. In https://cran.r-project.org/web/packages/emmeans/

Li, D., & Chen, L. (2022). Mitigating motion sickness in automated vehicles with vibration cue system. *Ergonomics*, *65*(10), 1313-1325.

Lin, Y., & Guo, G. (2022). *Research on the evaluation and relief with olfaction method of passengers carsickness degree* [master's thesis, Chongqing University].

Luo, R., Du, N., & Yang, X. J. (2022). Evaluating effects of enhanced autonomy transparency on trust, dependence, and human-autonomy team performance over time. *International Journal of Human–Computer Interaction*, *38*(18-20), 1962-1971.

Maculewicz, J., Larsson, P., & Fagerlönn, J. (2021). Intuitive and subtle motion-anticipatory auditory cues reduce motion sickness in self-driving cars. *International Journal of Human Factors and Ergonomics*, *8*(4), 370-392.

Manchon, J.-B., Bueno, M., & Navarro, J. (2022). How the initial level of trust in automated driving impacts drivers' behaviour and early trust construction. *Transportation Research Part F: Traffic Psychology and Behaviour*, *86*, 281-295.

Marc, A. (2015). *lm.br: Linear Model with Breakpoint*. In https://cran.r-project.org/web/packages/lm.br/index.html

Morimoto, A., Isu, N., Okumura, T., Araki, Y., Kawai, A., & Masui, F. (2008). Image rendering for reducing carsickness in watching onboard video display. In *2008 Digest of Technical Papers-International Conference on Consumer Electronics*.

Mosaferchi, S., Califano, R., & Naddeo, A. (2023). How personality, demographics, and technology affinity affect trust in autonomous vehicles: A case study. *Human Factors in Transportation*, *95*(95).

Nina. (2022). *Guangzhou allows driverless taxi service to operate in Nansha*. Newsgd. Retrieved October 5 from https://www.cnbayarea.org.cn/english/News/content/mpost_986771.html

Nolan, S. (2025). *Inside WeRide's global autonomous robotaxi expansion*. EV Magazine. Retrieved April from https://evmagazine.com/self-drive/weride-expands-autonomous-robotaxi-services-globally

Oliveira, L., Burns, C., Luton, J., Iyer, S., & Birrell, S. (2020). The influence of system transparency on trust: Evaluating interfaces in a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *72*, 280-296.

Qi, B., Guo, Q., & Liu, M. (2024). A study on the effects of different interaction modalities on driving trust in automated vehicles. In the *International Conference on Human-Computer Interaction*.

Reason, J. (1975). Motion sickness. In: Academic Press.

Rolnick, A., & Lubow, R. (1991). Why is the driver rarely motion sick? The role of controllability in motion sickness. *Ergonomics*, *34*(7), 867-879.

SAE, I. (2021). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles—SAE International. In.

Smith, C., Clegg, B. A., Heggestad, E. D., & Hopp-Levine, P. J. (2009). Interruption management: A comparison of auditory and tactile cues for both alerting and orienting. *International Journal of Human-Computer Studies*, *67*(9), 777-786.

Suwa, T., Sato, Y., & Wada, T. (2022). Reducing motion sickness when reading with head-mounted displays by using see-through background images. *Frontiers in Virtual Reality*, *3*, 910434.

Thomas, E., McCrudden, C., Wharton, Z., & Behera, A. (2020). Perception of autonomous vehicles by the modern society: A survey. *IET Intelligent Transport Systems*, *14*(10), 1228-1239.

Wang, Z., Hu, H., Sun, M., & He, D. (2024). An Exploration of Users’ Trust in and Willingness to Pay for Fully Driverless Vehicles After Their First Ride. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*.

Wenger, T. (2024). *I spent two days riding Waymo autonomous cars in Phoenix. It's perfect for travelers.* Matador Network. Retrieved May 1 from https://matadornetwork.com/read/waymo-phoenix/

Wickens, C. D., Dixon, S. R., & Seppelt, B. (2005). Auditory preemption versus multiple resources: Who wins in interruption management? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*.

Xie, W., He, D., & Wu, G. (2023). Inducers of motion sickness in vehicles: A systematic review of experimental evidence and meta-analysis. *Transportation Research Part F: Traffic Psychology and Behaviour*, *99*, 167-188.

Xie, W., Wang, Z., Jiang, Y., Huang, C., Wang, J., & He, D. (2025). Exploring the effects of regenerative braking and the auditory cues for alleviating motion sickness in electric vehicles. *International Journal of Human–Computer Interaction*.

Yang, X. J., Unhelkar, V. V., Li, K., & Shah, J. A. (2017). Evaluating effects of user experience and system transparency on trust in automation. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*.

Yunus, I., Jerrelind, J., & Drugge, L. (2021). Evaluation of motion sickness prediction models for autonomous driving. In *The IAVSD International Symposium on Dynamics of Vehicles on Roads and Tracks*.

Appendix

Table A.1. Trust questionnaire used before and after the ride

<table>
<tr><th>Question</th><th>Initial scale (before a ride)</th><th>Final scale (after a ride)</th></tr>
<tr><td>Q1</td><td>I would feel safe in the autonomous vehicle.</td><td>I felt safe in the autonomous vehicle.</td></tr>
<tr><td>Q2</td><td>The AV provides me with more safety compared to human drivers.</td><td>The AV provided me with more safety compared to human drivers.</td></tr>
<tr><td>Q3</td><td>I would trust the decisions of an AV in most situations.</td><td>I trusted the decisions of an AV in most situations.</td></tr>
<tr><td>Q4</td><td colspan="2">If the traffic conditions are complex, I will not take an AV.</td></tr>
<tr><td>Q5</td><td colspan="2">If the weather conditions are terrible (e.g., fog, glare, rain), I will not take an AV.</td></tr>
<tr><td>Q6</td><td>Rather than monitoring the traffic environment, I can focus on other activities confidently.</td><td>Rather than monitoring the traffic environment, I could focus on other activities confidently.</td></tr>
</table>