Transportation Research Part F: Psychology and Behaviour 122 (2026) 103760

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

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

Optimal lateral acceleration for different levels of automated driving: a test track study of passenger evaluation for curve negotiation

Chen Peng [a,b,*], Stefanie Carlowitz [a], Ruth Madigan [a], Claus Marberger [c], Michael Schulz [d], Sabine Osswalt [d], Andreas Schultz [d], Natasha Merat [a]

[a] *Institute for Transport Studies, University of Leeds, Leeds, United Kingdom.*
[b] *Transport Safety Research Centre, School of Design and Creative Arts, Loughborough University, United Kingdom*
[c] *Corporate Sector Research and Advanced Engineering, Robert Bosch GmbH, Renningen, Germany*
[d] *Cross-Domain Computing Solutions, Robert Bosch GmbH, Abstatt, Germany*

A R T I C L E I N F O

Keywords:
Automated driving
Lateral acceleration
Optimal range
Kinematics
User

A B S T R A C T

Lateral acceleration impacts passenger comfort in automated vehicles (AVs). This study investigated the optimal lateral acceleration range for AVs negotiating curves, considering curve radii, engagement in non-driving related activities (NDRAs), and automation level (SAE L1 vs L4). An experiment was conducted on a test track with 28 participants. Participants, seated in the driver's seat, experienced three lateral acceleration levels (2, 3, 4 m/s^2) across five different road curves. Evaluations, using a 7-point bipolar scale, were collected for both SAE Level 1 (assisted driving) and Level 4 (highly automated driving), with NDRA engagement considered under L4. Data were analysed using multilevel Bayesian multinomial logistic models.

Results suggest that no single universally "optimal" lateral acceleration was identified, as preferences were contextual. Generally, 2 m/s^2 was often rated "too slow", particularly without NDRAs in L4, or in L1 driving. Conversely, 4 m/s^2 was frequently perceived as "too fast". Within the tested range, the level of 3 m/s^2 often appeared to be a more balanced option, provided that road geometry, passenger activity, and automation level are accounted for. However, its optimality depended on other factors. Larger curve radii (gentler curves) led to a preference for lower lateral accelerations. NDRA engagement in L4 driving made 2 m/s^2 receive fewer "too slow" ratings and amplified the "too fast" perception for higher accelerations. L4 automation, compared to L1, also made 2 m/s^2 receive fewer "too slow" ratings. The findings provide detailed guidelines for designing AV lateral control systems, suggesting that lateral acceleration should be dynamically adjusted based on contextual factors to ensure comfort throughout the journey.

1. Introduction

Passenger comfort is essential for user acceptance and the uptake of Automated Vehicles (AVs). In automated driving, comfort can

* Corresponding author at: Institute for Transport Studies, University of Leeds, Leeds, United Kingdom.
E-mail addresses: c.peng@leeds.ac.uk, c.peng@lboro.ac.uk (C. Peng), tsseh@leeds.ac.uk (S. Carlowitz), r.madigan@leeds.ac.uk (R. Madigan), claus.marberger@de.bosch.com (C. Marberger), michael.schulz2@de.bosch.com (M. Schulz), Sabine.Osswalt@de.bosch.com (S. Osswalt), andreas.schultz@de.bosch.com (A. Schultz), n.merat@its.leeds.ac.uk (N. Merat).

https://doi.org/10.1016/j.trf.2026.103760
Received 2 October 2025; Received in revised form 26 July 2026; Accepted 27 July 2026
Available online 31 July 2026

be defined as a subjective state of ease and pleasantness, with the elimination of discomfort; it is shaped by a range of physical and psychological factors and fluctuates in response to user interaction with the AV and the surrounding driving environments (Peng, 2024).

Passenger comfort is found to be affected by the driving styles of AVs (Hartwich et al., 2018; Ossig et al., 2023; Peng et al., 2022; Yusof et al., 2016). An AV's driving style is affected by its kinematic behaviours - such as speed, acceleration, and deceleration – as well as its proximal behaviours - including spatial and temporal distance to other vehicles, and the negotiation of road geometries (Hajiseyedjavadi et al., 2022; Hartwich et al., 2018; Peng et al., 2024). Among these factors, acceleration (i.e., rate of change of speed) and jerk (i.e., rate of change of acceleration) have been highlighted as particularly important for passenger comfort in both traditional modes of passenger transport, such as buses and trains (Eboli et al., 2016; P. Eriksson & Friberg, 2000; Martin & Litwhiler, 2008) and AVs (Bae et al., 2019; de Winkel et al., 2023; Peng et al., 2025; Wang et al., 2020).

Vehicle motion occurs in six degrees of freedom: three translational (longitudinal, lateral, and vertical) and three rotational (roll, pitch, and yaw) (Turner & Griffin, 1999), and acceleration and jerk occur in all translational directions. While vertical motion is largely controlled through the suspension system and seat design, longitudinal and lateral motions directly relate to driving manoeuvres, such as braking, lane changing, and cornering. Notably, lateral acceleration, particularly during curves or roundabouts, has been identified as the primary cause of motion sickness for passengers in traditional vehicles (Turner & Griffin, 1999). Peng et al. (2025) also identified lateral acceleration as the most influential kinematic factor affecting passenger comfort during AV navigation through curved roads.

1.1. *Standards and thresholds for lateral acceleration*

Despite the importance of lateral acceleration, the optimal range of lateral acceleration for automated driving remains unclear in the literature. Bosetti et al. (2013) observed that active drivers rarely exceed lateral acceleration of $4\ m/s^2$, even for sharp turns and roundabouts (Bosetti et al., 2013). Eriksson and Svensson (2015) proposed that both longitudinal and lateral acceleration should be below $2\ m/s^2$ for comfort in AVs. However, such a threshold was conjectured based on research for buses and trains, and it was used to inform simulation and control system design, rather than being derived from passenger input. While there are no specific standards regarding the lateral acceleration of AVs negotiating curves, some ISO standards briefly address related aspects. The ISO 11270 standard for lane-keeping assistance system (LKAS) recommends that lateral acceleration resulting from lane-keeping actions should not exceed $3\ m/s^2$ (ISO, 2018), which may differ from appropriate values for curve negotiation. While the ISO 15622 standard for adaptive cruise control (ACC) systems focuses on detailed operational limits for longitudinal acceleration and deceleration, it briefly notes that average drivers tend to keep lateral acceleration within $2\ m/s^2$ when following vehicles on curves (ISO, 2018). Regardless, with this limited information, it is insufficient to inform the design of comfortable lateral acceleration for automated driving in a range of scenarios.

A recent study by de Winkel et al. (2023) investigated passenger discomfort in response to different acceleration and jerk profiles in automated driving, comparing their amplitude and direction. This study focused on *discrete* motion, and participants were asked to close their eyes to experience the motion without distraction. Sinusoidal and triangular pulses were used to simulate acceleration (peak value range 0.4–2 m/s^2) and jerk (peak value range 0.5–15 m/s^3); the former are smooth and continuous sine waves, simulating more gradual changes in acceleration, while the latter are linear waves, representing more rapid and abrupt changes in acceleration. Participants rated their discomfort after each motion pulse, expressing how uncomfortable they felt by scaling their ratings relative to a baseline with a value of 100. For example, if a motion felt twice as uncomfortable, it should be rated as 200. The numerical ratings were then mapped by participants to seven-point verbal 'quantifiers', including excellent, very good, good, so-so, bad, very bad, and terrible. For lateral acceleration, the authors found that the average 'so-so' point was $0.98\ m/s^2$, which was considered a threshold for acceptable acceleration. However, this threshold appears to be much lower than those observed during active driving or driving with assistance systems, which warrants further investigation. More importantly, driving on real roads is always about responding to other road users, objects, and road geometries, which highlights the need for research going beyond solely experiencing the discrete motion under eye-closed conditions.

1.2. *Contextual influences on passenger preference*

A comfortable driving style of AVs is expected to accommodate varied road environments, including curves and roadside elements (Carlowitz, Madigan, et al., 2024; Hajiseyedjavadi et al., 2022; Stayton & Stilgoe, 2020). As changes in vehicle dynamics result from responses to road environments, the effect of curve radii on lateral acceleration choices is noticeable. In general, passengers' comfort has been shown to decrease when AVs navigate sharp curves (i.e., curves with smaller curve radii) at high speeds (Hajiseyedjavadi et al., 2022; Peng et al., 2022). However, few studies have explicitly examined the impact of curve radii on passenger perception and preference of specific vehicle kinematics.

The potential effect of the level of automation on passengers' perceived optimal lateral acceleration remains underexplored (Beggiato et al., 2015; Madigan et al., 2018). Specifically, when lower levels of driving assistance or automation are engaged (SAE Level 3 and below), users mostly act as active drivers and/or system supervisors, maintaining awareness of the changing vehicle dynamics in response to external environments. In contrast, at higher levels of automation (SAE Level 4 and above), users primarily take on passive roles, without control over the vehicle and with limited ability to anticipate upcoming manoeuvres (Elbanhawi et al., 2015; SAE, 2021). Active drivers and passive passengers are likely to have different preferences for driving styles (see evidence in Miller et al., 2014; Mühl et al., 2019; Peng et al., 2022; Price et al., 2016). For example, Price et al. (2016) found that precise lane centring increased drivers' trust when using lane-keeping assistance systems, while Peng et al. (2025) found that passengers in highly automated driving tended to prefer curve-cutting behaviours. However, research that directly compares user preferences for driving styles in different levels of automation is rare. Future AVs may dynamically switch between automation levels during a ride (Kyriakidis et al., 2019; Tinga et al., 2023), which necessitates reassessment of comfortable vehicle control strategies tailored to different levels of automated driving. Moreover, passengers are allowed to engage in non-driving-related activities (NDRAs) in highly automated driving, which can increase the possibility of motion sickness (Brietzke et al., 2021; Kato & Kitazaki, 2006; Metzulat et al., 2024). It adds more uncertainty to the understanding of the optimal range of lateral acceleration.

1.3. The present study

In the present study, we investigate passengers' evaluation of lateral acceleration as an AV negotiates curves on a test track, aiming to establish an optimal zone for lateral acceleration. The effect of engagement in NDRAs is considered for highly automated driving. Moreover, to understand the effect of the level of automation, we also compare passengers' evaluation for lateral acceleration across two automation levels: assisted (SAE Level 1) and high (SAE Level 4). Regarding the Level 4 system, the vehicle operated in fully autonomous mode within a specific area and the driver had no responsibility for the driving task. Thus, the aim was to investigate optimal lateral acceleration in contexts where the user was quite engaged in the driving task (L1), and where they had no responsibility for the driving task (L4).

We aim to address the following research question:

What are the optimal lateral acceleration values of AVs in curve driving?

Three levels of lateral acceleration were implemented for this study: 2, 3, and 4 m/s^2 (see more details in section 2.3). The following sub-questions are also considered:

1) What impact do different curve radii have on preferences for lateral acceleration?
2) What impact does the NDRA engagement have on preferences for lateral acceleration?
3) What impact does the automation level (SAE L4, SAE L1) have on preferences for lateral acceleration?

2. Method

The experiment used a within-participant design, where all participants experienced all combinations of lateral acceleration (3 levels: 2, 3, and 4 m/s^2), road curves (5 curves), and automation levels (2 levels: SAE Level 1 and Level 4) across two rides. The high automation level (SAE Level 4) also included an additional factor to compare riding with and without NDRAs. Participants were seated in the driver's seat in the test vehicle, with the safety driver sitting in the co-driver seat. Participants were asked to provide ratings for the cornering speed after the test vehicle finished each cornering manoeuvre. Statistical models were fitted to the data to determine the

Table 1
A summary of demographic characteristics of the sample.

Item		All	Female	Male
		Number of participants		
Driving experience per year				
	5.000 km - 15.000 km	10	6	4
	15.000 km - 30.000 km	14	6	8
	More than 30.000 km	3	0	3
Preference for being driven at co-driver's seat				
"Do you like to be driven in the co-driver's seat?"				
	No	8	1	7
	Yes	19	11	8
		Average sum across participants		
Desired manual driving style				
"How do your friends rate your manual driving style?" 0 – less dynamic to 100 – more dynamic				
	Mean (SD)	62.33 (25.78)	57.83 (26.90)	65.93 (25.18)
Desired automated driving style				
"How would you like to be driven by an automated vehicle?" 0 – less dynamic to 100 – more dynamic				
	Mean (SD)	54.95 (24.04)	50.92 (28.30)	58.13 (20.48)
Affinity for cars				
1 – "I find cars boring" to 7 – "cars excite me"				
	Mean (SD)	5.78 (1.55)	5.17 (1.59)	6.27 (1.39)
Affinity for technology				
1 – "I find technology boring" to 7 – "Technology excite me"				
	Mean (SD)	5.63 (1.33)	4.75 (1.48)	6.33 (0.62)

* One participant's demographic data was missing. Content in *italic* represents the specific question asked, and the corresponding response scale.

preference for different lateral accelerations in a range of conditions.

2.1. Participants

Twenty-eight participants (12 female, 15 male, 1 unspecified; mean age = 47.0, SD = 13.7) completed the experiment. Individuals susceptible to motion sickness were excluded from the recruitment process. Table 1 provides a summary of the sample's demographic characteristics and other information. Overall, participants reported having moderately dynamic manual driving styles and expected similar levels of dynamic features in automated driving styles. Their interest in technology and cars was also found to be moderate.

Participants were recruited by an independent external research institute and were compensated for their participation, as well as reimbursement for travel expenses to the test track. Ethical approval was obtained from an independent ethics committee of the RUMBA Project Consortium (protocol name “Bosch2023c”).

2.2. The vehicle and the test track

The vehicle used for this study was a VW Golf 7 Variant, equipped with an automated system capable of controlling both longitudinal and lateral driving operations, aiming to represent a Level 4 automation feature (SAE, 2021). To ensure consistent vehicle movements, a vehicle-in-the-loop system was employed, with which the vehicle's movements were based on a pre-recorded GPS route, rather than real-time road features (e.g., lane markings) (Fig. 1).

Participants were seated in the driver's seat, while a safety driver occupied the co-driver's seat to intervene if necessary. The experimenter sat in the back seat to collect immediate feedback from participants after each scenario. No other vehicles or vulnerable road users were present on the test track during the study.

The experiment took place on a closed test track in Bad Sobernheim, Germany, managed by TRIWO Automotive Testing GmbH. The

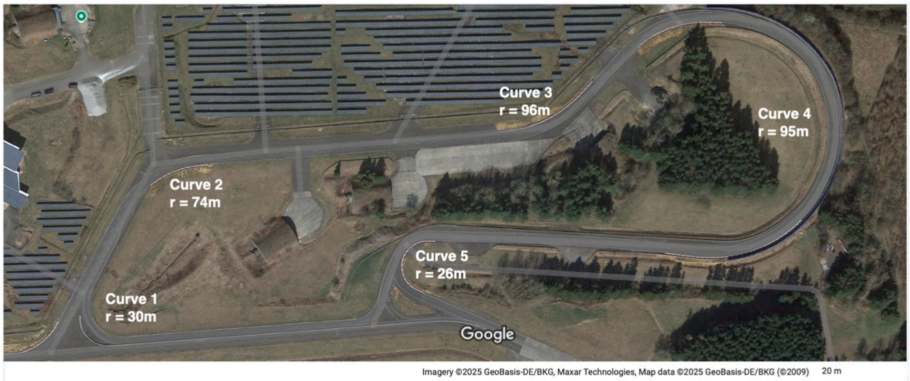

Fig. 1. The vehicle (top), and the test track (bottom) used for the experiment, featuring a variety of curves. Note that the curve following Curve 5 was excluded from the experimental design for both technical and practical reasons. The specific lengths of curves (C) and the connecting straight segments: **C1:** 65 m (Segment to C2: 75 m); **C2:** 75 m (Segment to C3: 231 m); **C3:** 231 m (Segment to C4: 44 m); **C4:** 407 m (Segment to C5: 186 m); **C5:** 51 m (Segment to C1: 189 m). The picture of the test track was captured from Google Maps, and additional annotations for curve radii were added.

total length of the test track is approximately 1450 m. Each curved lap took around 1.5 min to complete, and the entire study lasted about one hour.

2.3. Lateral acceleration

In this experiment, the AV navigated various curves at different speeds, with the aim of controlling lateral acceleration (a_y), which acts perpendicular to the vehicle's direction of travel. According to the *Centripetal Acceleration formula,* $a_y = v^2/r$, where v is vehicle speed and r is curve radius. By adjusting the vehicle's speed, the peak lateral acceleration could be controlled. Specifically, for a given curve radius, lateral acceleration increases with speed. Conversely, maintaining a constant lateral acceleration across curves with different radii requires higher speeds on curves with larger radii.

Three peak lateral acceleration values were implemented in this study: 2 m/s^2, 3 m/s^2, and 4 m/s^2 (Fig. 2). These values were selected to represent a range of typical lateral acceleration levels that a vehicle might experience in real-world driving, according to thresholds suggested by previous studies (Bosetti et al., 2013; P. Eriksson & Friberg, 2000; see also details in section 1.1). Moreover, according to Bae et al. (2022), such values correspond to behaviours of 'normal drivers', avoiding both overly cautious and excessively aggressive driving styles. It is worth noting that the threshold of *0.98 m/s^2* for discrete lateral motions proposed by de Winkel et al. (2023) was excluded, as pilot testing suggested it to be too slow for naturalistic driving conditions. Due to the fixed geometric layout of the test track, changes in lateral acceleration were inherently associated with variations in jerk (the rate of change of acceleration). While these factors are physically linked during curve negotiation, the experimental design focused on target lateral acceleration levels to establish comfort thresholds.

2.4. Subjective evaluation

During the Level 4 drives, participants were instructed to imagine that the AV could drive them autonomously along a familiar route – such as commuting to work - while they could engage in other activities. The AV could offer different driving profiles (e.g., Sport, Normal, Eco), each determining the maximum speed in curves. In the present study, participants were asked to focus on the "Normal" profile and evaluate the speeds in various curves, judging whether they were too fast, too slow, or appropriate for their personal interpretation of an optimal "Normal" driving profile. Participants were prompted to provide feedback for the curve speeds in the last third of each curve, signalled by an audible cue triggered automatically by GPS.

During the Level 1 drives, participants were asked to imagine that their vehicle was equipped with an automatic speed controller

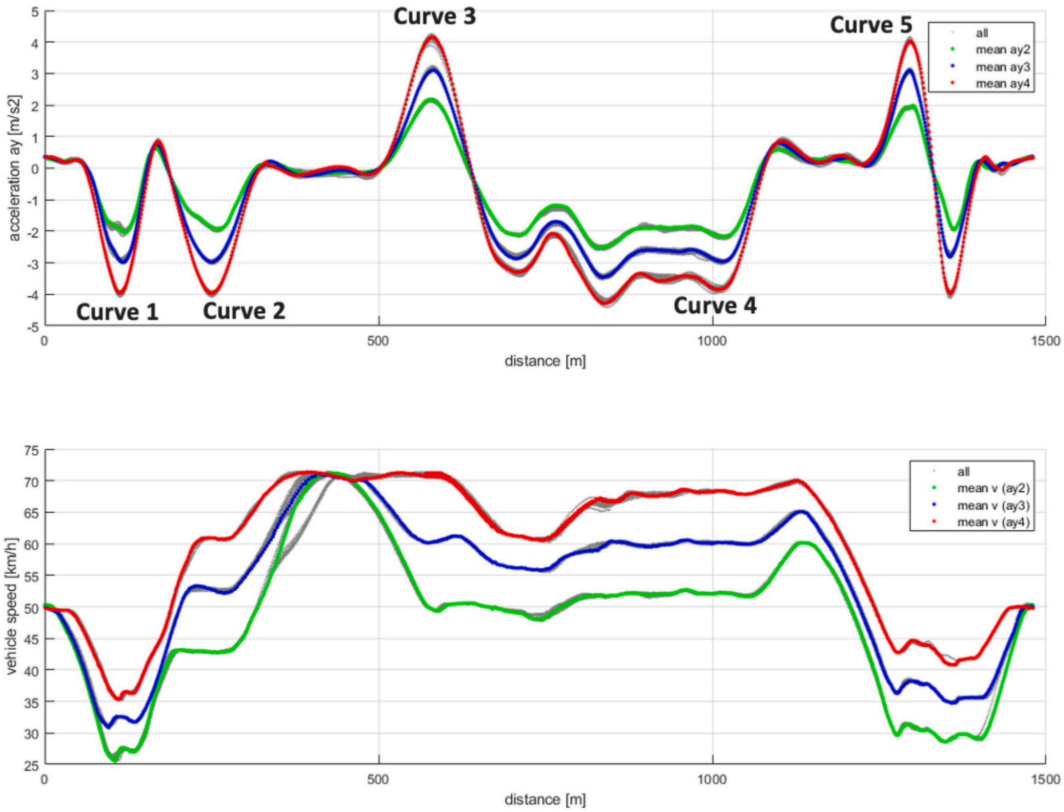

Fig. 2. Lateral acceleration (top) and vehicle speed (bottom) along the route.

that adjusts the speed based on the road conditions. They were instructed to maintain the set speed while taking over the steering to control the vehicle's direction through the curves. As in the Level 4 drives, participants were prompted by an audible signal in the last third of each curve to provide feedback on the curve speed. In the NDRA condition, they were instructed by the experimenter to immediately return to the NDRA following their rating. While individual steering behaviour could cause minor deviations in the actual acceleration experienced, our analysis uses the pre-set target levels as predictors. This aligns with our research focus on understanding users' perception and evaluation of the vehicle's intended kinematic behaviours in different driving contexts.

For providing evaluation, participants were instructed to answer the question: “How would you rate the cornering speed you just experienced?” They were asked to quickly and intuitively choose a value on the provided scale (Fig. 3) and verbalise their response, which was then recorded by the experimenter. The question on cornering speed was used because a pre-study test showed participants understood it better than perceived lateral forces or acceleration levels.

This bipolar scale, sometimes called the Just-About-Right (JAR) scale, is regularly used in consumer research to evaluate if a product (e.g., yoghurt) is at its optimal level or at too-high or too-low level. It is often combined with another overall liking rating for analysis (Paries et al., 2022). For practical consideration, we did not include another overall liking rating. The rating scale used in this study includes both directional information (fast or slow) and its intensity (how fast or slow the AV is perceived to be), with the middle point indicating the optimal option. According to Rothman and Parker (2009), JAR scales typically range from three to nine points. While a three-point scale is the theoretical minimum, using a wider scale for data collection helps mitigate “end avoidance”, allowing participants to express more detailed opinions without feeling forced towards the centre.

2.5. Procedure

Prior to the experiment day, participants completed an online survey, in which they provided demographic information, reported personal characteristics (e.g., interest in cars and sensation seeking), and gave informed consent to participate in the study. As shown in Fig. 4, on the experiment day, participants received instructions and were introduced to the scale they would use for evaluation during the drives. Participants were driven to the test track by the safety driver, including one introductory round to demonstrate the track.

Depending on group assignment, participants started the experiment with either the SAE Level 0 (manual driving), Level 1 (L1), or Level 4 (L4) condition. For manual driving, participants drove the test track for three rounds. While the primary purpose of the L0 condition was to characterise participants' manual driving styles for a baseline comparison, these data are not presented in the current paper as they fall outside of the scope. For L1 and L4, each condition was preceded by two additional drives, to allow participants to familiarise themselves with the respective automation system and understand the timing of evaluation prompts.

For the L1 trials, participants were responsible for steering the vehicle and completed six randomised laps with lateral acceleration values of 2, 3, and 4 m/s^2 (each condition repeated twice). For the L4 trials, participants went through 12 randomised laps at the same lateral acceleration levels (2, 3, and 4 m/s^2), both with and without an NDRA, with each condition repeated twice. The NDRA condition was alternated to reduce potential motion sickness. For the drives with an NDRA, participants were asked to play a smartphone-based game, in which they tapped the largest circle shown on the screen (https://apkpure.com/surt-mobile/de.lapoehn.surt). This

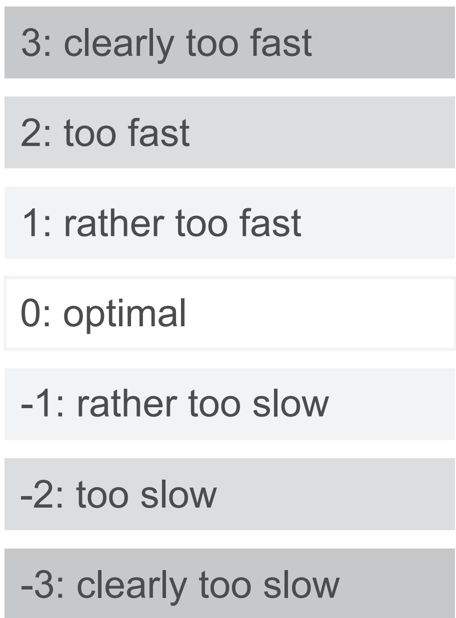

Fig. 3. The scale used by participants for evaluation.

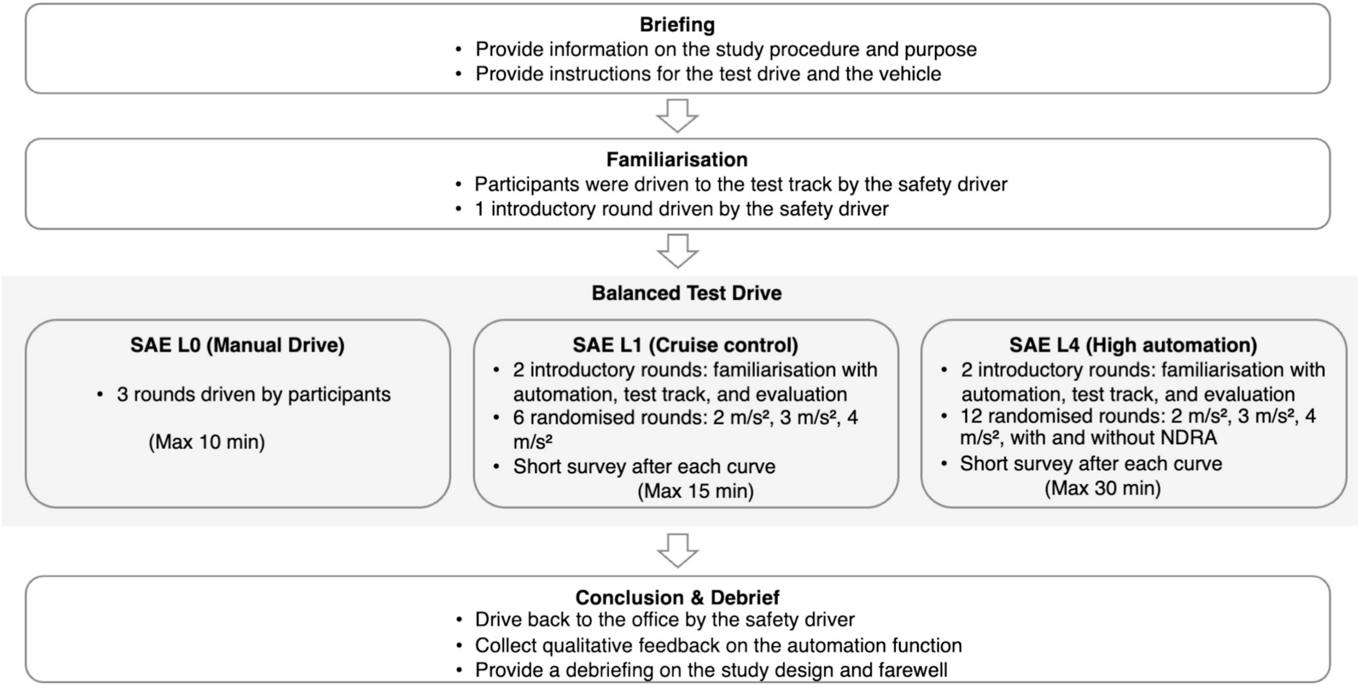

Fig. 4. Procedure of the experiment.

NDRA is adapted from the Surrogate Reference Task (SuRT; *ISO 17488:2016*), a standardised visual-manual task frequently used as non-driving related tasks in driving research (e.g., Guo et al., 2020; Young et al., 2013). The game application recorded their performance, awarding higher scores for faster and more accurate taps. Participants were encouraged to achieve the highest possible score. Each experimental condition was repeated twice to ensure the reliability of subjective evaluations and to mitigate potential order effects.

2.6. Statistical analysis

A *multilevel multinomial logistic model under the Bayesian framework* was adopted for the statistical analysis. The model predicts the probability of observing each of the rating categories as a function of the predictor variables.

Multilevel models (or mixed-effects models) are suitable for data with hierarchical or clustered structures; for example, repeated ratings from the same participant might be correlated with each other. Multilevel models include fixed effects (i.e., effects that remain constant across observations) and random effects (i.e., variability and differences between participants) (Goldstein, 2011).

To reduce the complexity resulting from the bipolar scale, we re-coded the 7-point scale to a 3-point scale, by combining all negative ratings (-3,-2,-1) and all positive ratings ($+1$, $+2$, $+3$), respectively. The re-coded scale includes three levels: Too Slow, Optimal, and Too Fast. This recode is conducted for two reasons: a) combining ratings can increase the number of data points in each category, which enhances the model fit, and b) participants tended not to select extreme ratings, resulting in fewer data points for these categories. In addition, we tested the original 7-point data and an alternative way of combining categories. However, the 7-point model added too much complexity and had difficulty converging, while the alternative grouping tended to mask important differences between categories. We decided to treat the scale as nominal and set the middle point as the reference category. In this way, we were able to understand whether the rating deviated towards the fast or the slow side (i.e., whether the rating tends to be above or below zero) compared with an optimal rating. For this purpose, the *multinomial* logistic *model* was chosen (Agresti, 2003; Hedeker, 2003).

Bayesian statistics is an approach to data analysis and parameter estimation based on Bayes' theorem, where prior knowledge about parameters is updated with the information obtained from observed data. The posterior distribution reflects the updated knowledge, combining prior knowledge with observed data (Bayes, 1763). Moreover, the Bayesian inference has flexibility in random effects modelling for multinomial models, the distribution of the data, and the sample size (van de Schoot et al., 2021). The 95% credible interval of posterior distributions indicates that the interval has a 0.95 probability of containing the value of the parameter.

For model performance comparisons, *leave-one-out* (*LOO*) *cross-validation* was used. LOO is a method for measuring the predictive accuracy of a fitted Bayesian model using the log-likelihood evaluated at the posterior simulations of the parameter values (Vehtari et al., 2017).

The R package *brms* was used for the analysis (Bürkner, 2017, 2017; Heiss, 2023). The scripts can be found at https://github.com/pchencecilia/comfort_acc.

2.6.1. Model description

Let Y_{ki} and $\mu_{k,i}$ be the choice of ratings and the probability of rating Optimal ($k = 0$), Too Slow ($k = 1$), Too Fast ($k = 2$), for participant i. We assume that Y_{ki} follows a multinomial distribution: $Y_{ki} \sim Multinomial(1, \boldsymbol{\mu}_i)$, where $\boldsymbol{\mu}_i = (\mu_{0,i}, \mu_{1,i}, \mu_{2,i})$ is the vector of probabilities.

The probability of selecting rating category k relative to the reference category $k = 0$ (optimal) is determined by the linear predictor, $\eta_{k,i}$:

$$\mu_{k,i} = \frac{\exp(\eta_{k,i})}{1 + \sum_{l=1}^{2} \exp(\eta_{l,i})}, k = 1, 2$$

To identify the model, the linear predictor for the reference category is set to zero, which is $\eta_{0,i} = 0$. Then, the probability for selecting the reference category $k = 0$ (is)

$$\mu_{0,i} = \frac{1}{1 + \sum_{l=1}^{2} \exp(\eta_{l,i})}$$

The log-odds of rating category k relative to the reference category are modelled as:

$$\ln\left(\frac{\mu_{k,i}}{\mu_{0,i}}\right) = \eta_{k,i} = \boldsymbol{\beta}_k \boldsymbol{X}_i + u_{ki}, k = 1, 2,$$

where $\boldsymbol{X}_i$ is the row vector of independent variables for participant i, and $\boldsymbol{\beta}_k$ is the vector of regression coefficients corresponding to the independent variables for the log-odds of rating category k versus rating category 0. $u_{ki} \sim N(0, \sigma_k^2)$ represents the random intercept for participant i.

The linear predictor $\boldsymbol{\beta}_k\boldsymbol{X}_i$ can be written explicitly as:

$$\beta_k X_i = \beta_{k0} + \beta_{k1}X_{1i} + \beta_{k2}X_{2i} + \ldots + \beta_{kp}X_{pi}$$

where $\beta_{k0}$is the intercept, X_{pi} is the value of the p-th independent variable for participant i, and β_{kpi} is the corresponding coefficient.

We used weakly informative priors $\beta_k \sim N(0,1)$ for fixed effect parameters and $\sigma_k \sim HalfNormal(0,1)$ for between-participant variability.

For model result interpretation, we used the odds ratio rather than the direct output in logit to explain the influence of the variables on ratings. The odds ratio is the exponential transformation of the estimates (OR $= e^{estimate}$): a value of 1 means no effect, while values greater than 1 suggest that a one-unit increase in the studied variable increases the odds of the rating by $\left(\left(e^{estimate} - 1\right) \times 100\%\right)$.Conversely, a value less than 1 means the odds decrease by $\left(\left(1 - e^{estimate}\right) \times 100\%\right)$. A variable is considered *significant* or *meaningful* if its credible interval (CI) in odds ratio **excludes 1** (see examples using the odds ratio for interpretations: (Chen et al., 2016; Delbiso et al., 2016; Gudmestad & Metzger, 2025).

2.6.2. Consistency of repeated ratings

To assess the consistency of subjective evaluation across the two repetitions, ratings from the first and second exposures to the same condition were compared per participant, across lateral acceleration levels, curve radii, and SAE automation levels. A Spearman correlation (Rho $= 0.84$, $p < .001$) indicated significant correlation between the two ratings. Moreover, the Intraclass Correlation Coefficient (ICC3 $= 0.85$) confirmed high test-retest reliability of the subjective evaluations.

3. Results

3.1. Evaluation of lateral acceleration in L4 automated driving

A mixed effects model was fitted to examine the effects of lateral acceleration levels, engagement in NDRAs, curve radii, and the ride order on participants' perception and evaluation of lateral acceleration, as measured on the recoded 3-point scale (see Formula (1)).

Table 2

Posterior estimation results based on the multilevel Bayesian logistic model (*Rating ~ ay + NDRA + Curve + Order + ay:NDRA + ay:Curve + ay:Order + NDRA:Curve + NDRA:Order + Curve:Order + (1 | participant)*).

Parameter	Est.	SE	Est (OR)	CI.low	CI.high	Est.	SE	Est (OR)	CI.low	CI.high
Fixed effects										
Too slow (vs optimal)						*Too fast (vs optimal)*				
(Intercept)	1.30	0.36	3.66	1.81	7.33	−4.41	0.55	0.01	0.00	0.03
ay (reference: ay[2])										
ay[3]	−2.89	0.31	**0.06**	**0.03**	**0.10**	1.31	0.42	**3.71**	**1.67**	**8.63**
ay[4]	−5.05	0.48	**0.01**	**0.00**	**0.02**	4.97	0.44	**143.44**	**61.80**	**347.14**
NDRA (reference: NDRA[0])										
NDRA[1]	−2.31	0.27	**0.10**	**0.06**	**0.17**	−0.13	0.42	0.88	0.39	2.00
Curve radius	−2.45	0.34	**0.09**	**0.04**	**0.17**	1.12	0.48	**3.08**	**1.21**	**7.76**
Order (reference: Order[1])										
Order[2]	0.65	0.27	**1.91**	**1.10**	**3.24**	−0.73	0.42	0.48	0.21	1.09
ay × NDRA (reference: ay[2] × NDRA[0]) [a]										
ay[3] × NDRA[1]	0.30	0.38	1.35	0.63	2.84	0.99	0.46	**2.70**	**1.08**	**6.71**
ay[4] × NDRA[1]	0.40	0.61	1.48	0.44	4.61	0.71	0.43	2.04	0.88	4.71
ay × Curve radius (reference: ay[2] × Curve radius) [a]										
ay[3] × Curve radius	−0.89	0.41	**0.41**	**0.19**	**0.89**	1.13	0.50	**3.11**	**1.15**	**8.27**
ay[4] × Curve radius	0.16	0.60	1.18	0.35	3.70	1.25	0.47	**3.50**	**1.37**	**8.65**
ay × Order (reference: ay[2] × Order[1]) [a]										
ay[3] × Order[2]	−1.58	0.39	**0.21**	**0.09**	**0.44**	0.07	0.46	1.07	0.43	2.60
ay[4] × Order[2]	−1.17	0.53	**0.31**	**0.11**	**0.87**	0.34	0.43	1.41	0.61	3.26
NDRA[1] × Curve radius	0.77	0.36	**2.17**	**1.03**	**4.37**	−0.79	0.35	**0.46**	**0.23**	**0.91**
NDRA[1] × Order[2]	0.33	0.33	1.39	0.73	2.59	0.54	0.33	1.71	0.89	3.33
Curve radius × Order[2]	−0.34	0.35	0.71	0.35	1.41	0.62	0.35	1.87	0.94	3.76
Random effects										
sd(u_1)	1.54	0.23								
sd(u_2)	1.93	0.27								
cor(u_{12})	−0.70	0.11								

Note. Estimate and SE are in log-odds, while Est (OR) and Credible intervals (CI) are in odds ratio. Bold values indicate significance. Curve radius in the model was standardised. For random effects, sd(u_1) and sd(u_2) represent the standard deviation of the random intercepts for "Too Slow (vs Optimal)" outcome category and "Too fast (vs Optimal)" category across individuals (the grouping variable), respectively. The term cor(u_{12}) represents the correlation between the two random intercepts across participants. [a] Interaction terms represent the change in predictor effects at ay = 3 m/s^2 or ay = 4 m/s^2 relative to the baseline combination of ay = 2 m/s^2 and the respective reference level (e.g., Order = 1, NDRA = 0).

$$Rating \sim ay + NDRA + Curve + Order + ay : NDRA + ay : Curve + ay : Order + NDRA : Curve + NDRA : Order + Curve : Order + (1 \mid participant)) \tag{1}$$

This section only includes conditions under SAE Level 4 driving. The distribution of the data and a comparison of alternative model specifications are provided in Appendix 1.

3.1.1. Model result

Table 2 presents the posterior estimation results from the random intercept Bayesian multilevel multinomial logistic regression model, which compares participants' ratings for conditions experienced under L4 automated driving. In this model, coefficients represent the relationship between each predictor variable and the log-odds of relative rating selection, while odds ratios (ORs) quantify the relative odds of that rating selection. As the "Optimal" rating category was set as the reference level, the effect estimates are interpreted in relation to the odds of receiving an "Optimal" rating (i.e., Too Slow vs Optimal, or Too Fast vs Optimal) (Agresti, 2003; Gudmestad & Metzger, 2025).

Fig. 5 complements these results by showing the predicted probabilities, as percentages, for each rating category across different conditions. These probabilities were derived using the softmax transformation of the model's linear predictors.

Fixed effects.

Evaluation of lateral acceleration levels.

Before examining changes due to higher lateral accelerations or other factors, the *baseline* condition is considered: at *a lateral acceleration of 2 m/s²*, *without NDRA engagement*, for an *average standardised curve radius*, and during *the first ride*. The odds of rating the condition as Too Slow (vs Optimal) were 3.66 times higher (OR = 3.66, CI: 1.81–7.33), and the odds of rating the condition as Too Fast (vs Optimal) were very low (OR = 0.01, CI: 0.00–0.03). This indicates that this baseline condition of 2 m/s^2 had a tendency to be considered as Too Slow rather than Optimal and had a very low tendency to be considered as Too Fast.

Compared to the lateral acceleration of 2 m/s^2, 3 m/s^2 was associated with 94% lower odds of being rated as Too Slow (vs Optimal; OR = 0.06, CI: 0.03–0.10). Conversely, the odds of being rated as Too Fast (vs Optimal) at 3 m/s^2 were 271% higher than at 2 m/s^2 (OR = 3.71, CI: 1.67–8.63).

At a lateral acceleration of 4 m/s^2 compared to 2 m/s^2, the odds of Too Slow ratings (vs Optimal) were substantially lower (OR = 0.01, CI: 0.00–0.02). The odds of Too Fast ratings (vs Optimal) at 4 m/s^2 were markedly higher (OR = 143.44, CI: 61.80–347.14).

Visual inspection of the predicted probabilities (Fig. 5) suggests that the lateral acceleration of 3 m/s^2 frequently resulted in the highest probability of Optimal ratings across various conditions; however, this optimal choice could shift based on other interacting factors.

The effect of NDRA engagement.

At a lateral acceleration of 2 m/s^2, engagement in the visual NDRA (compared to no engagement) significantly reduced the odds of Too Slow ratings (vs Optimal) by 90% (OR = 0.10, CI: 0.06–0.17). Also at 2 m/s^2, the main effect of NDRA engagement on the odds of Too Fast ratings (vs Optimal) was not statistically significant (OR = 0.88, CI: 0.39–2.00).

Regarding the interaction effect, the effect of NDRA engagement (vs no engagement) on the odds of Too Fast ratings (vs Optimal) was significantly stronger at lateral accelerations of 3 m/s^2 (OR = 2.70, CI: 1.08–6.71), compared to its effect at 2 m/s^2. This indicates that NDRA engagement amplified the tendency to rate lateral acceleration of 3 m/s^2 as Too Fast. While the same trend was observed for lateral acceleration of 4 m/s^2, the effect was non-significant (OR = 2.04, CI: 0.88–4.71). This might reflect a ceiling effect as 4 m/s^2 was already frequently rated as Too Fast regardless of NDRA engagement.

The effect of curve radii.

At a lateral acceleration of 2 m/s^2, a one standard deviation increase in curve radius (i.e., gentler curves) significantly decreased the odds of Too Slow ratings (vs Optimal) by 91% (OR = 0.09, CI: 0.04–0.17). Conversely, it increased the odds of Too Fast ratings (vs Optimal) by 208% (OR = 3.08, CI: 1.21–7.76).

Regarding interaction effects, at a lateral acceleration of 3 m/s^2, a one standard deviation increase in curve radius further reduced the odds of Too Slow ratings (vs Optimal) by 59% (OR = 0.41, CI: 0.19–0.89), compared to the effect of curve radius at 2 m/s^2. The positive effect of larger curve radii on the odds of Too Fast ratings (vs Optimal) was significantly stronger at 3 m/s^2 (OR = 3.11, CI: 1.15–8.27) and 4 m/s^2 (OR = 3.50, CI: 1.37–8.65), compared to its effect at 2 m/s^2.

There was a significant interaction between NDRA engagement and curve radius. Results showed that as curve radius increases (and correspondingly vehicle speed increases), the probability of Too Slow ratings decreases in both NDRA engagement and non-engagement conditions. However, the reduction is less pronounced when passengers are engaged in NDRAs (OR = 2.17, CI: 1.03–4.37), suggesting NDRA engagement weakens the moderating effect of curve radius on Too Slow ratings. Similarly, the reduction in Too Fast ratings with increasing curve radius is also weakened under NDRA engagement (OR = 0.46, CI: 0.23–0.91).

The effect of order.

At a lateral acceleration of 2 m/s^2, compared with the first ride, the odds of providing Too Slow ratings (vs Optimal) in the second ride significantly increased by 91% (OR = 1.91, CI: 1.10–3.24).

Regarding interaction effects, the tendency to rate higher lateral accelerations as Too Slow (vs Optimal) during the second ride (compared to the first ride) was reduced relative to this effect at 2 m/s^2. Specifically, the odds ratio for the second ride was 0.21 times

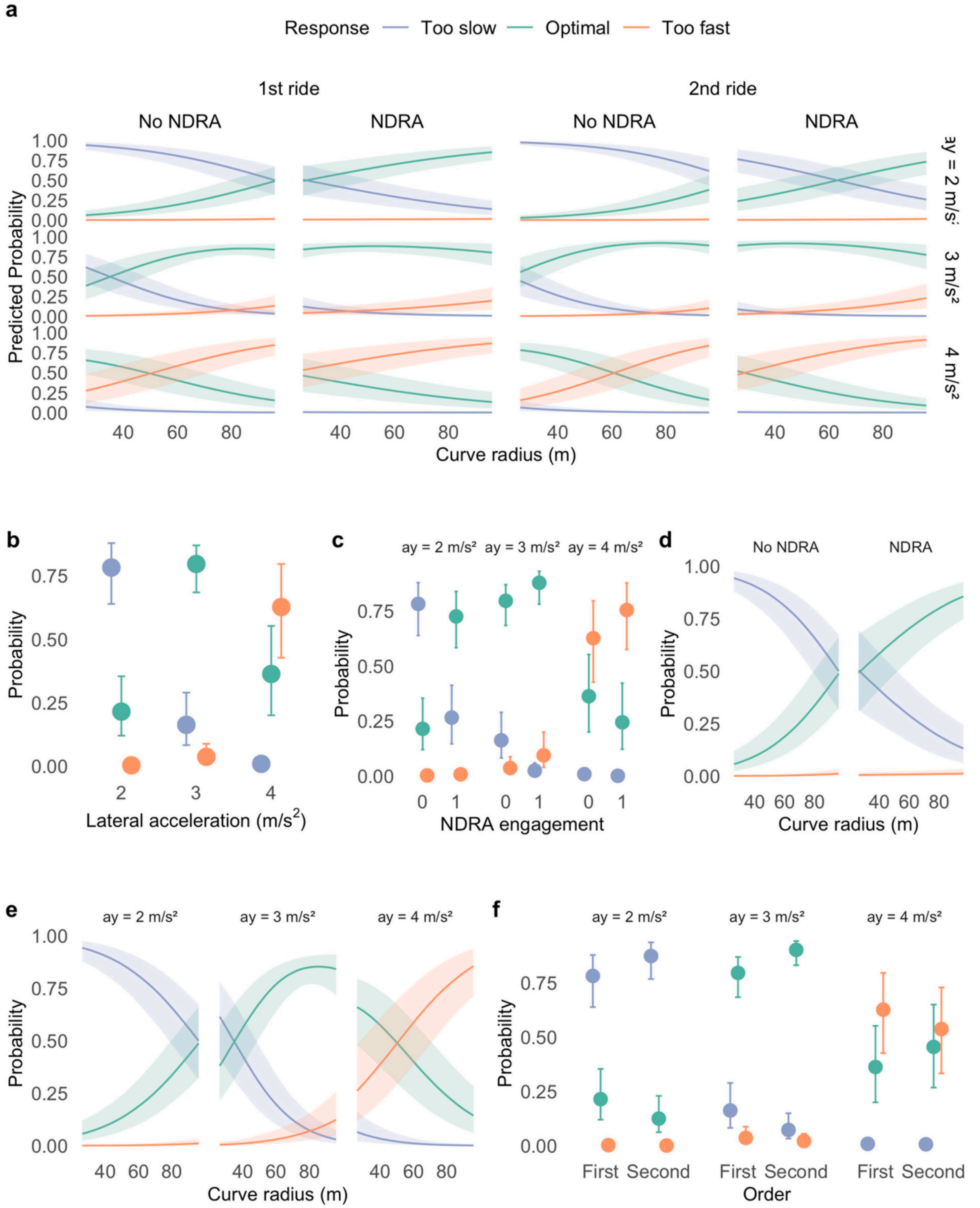

Fig. 5. Predicted probability of each rating category. (a) Predicted probabilities conditioned by lateral acceleration, NDRA engagement, curve radius, and ride order (b) Predicted probabilities by lateral acceleration. (c) Predicted probabilities by lateral acceleration and NDRA engagement. (d) Predicted probabilities by NDRA engagement and curve radius. (e) Predicted probabilities by lateral acceleration and curve radius. (f) Predicted probabilities by lateral acceleration and ride order. For plots (b)-(f), predictions were generated holding unplotted categorical predictors at their reference level and continuous predictors at their mean. For all plots, the probabilities for the three rating categories sum to 1. Error bars represent 95% CI.

smaller at 3 m/s^2 (OR = 0.21, CI: 0.09–0.44) and 0.31 times smaller at 4 m/s^2 (OR = 0.31, CI: 0.11–0.87), when predicting Too Slow ratings, compared to its effect at 2 m/s^2. The interaction effects of ride order with NDRA engagement and with curve radius were both found to be non-significant.

Radom effects.

As shown in Table 2, the random effects estimates indicate substantial between-participant variability in rating tendencies. The strong negative correlation (−0.70) between these random intercepts suggests that participants who are more likely to rate a condition Too slow (vs Optimal) are substantially less likely to rate it Too fast (vs Optimal), and vice-versa, independent of the fixed predictors.

3.2. Evaluation of lateral acceleration across automation levels

A mixed effects model was fitted to examine the effects of lateral acceleration levels, automation level, and curve radii on participants' evaluation of lateral acceleration, measured using the recoded 3-point rating scale (see Formula (2)).

$$Rating \sim ay + Level + Curve + ay : Level + ay : Curve + (1 \mid participant)) \quad (2)$$

To ensure comparability across automation levels, this analysis specifically excluded NDRA engagement conditions under Level 4 driving. The distribution of the data and a comparison of alternative model specifications are provided in Appendix 2. In particular, the interaction effect of automation level and curve radius was examined in sensitivity analysis; it was found to be non-significant and did not affect other effects. LOO test showed no significant improvement (see Appendix 2). Therefore, this interaction effect was not included in the model.

3.2.1. Model result

Table 3 details the posterior estimation results from the random intercept Bayesian multilevel multinomial logistic regression model, which compares participants' ratings of conditions experienced under L1 and L4 automated driving. To further illustrate these findings, Fig. 6 visualises the predicted probabilities (as percentages) for each rating category across different conditions.

Fixed effects.

Evaluation of lateral acceleration.

The evaluation of lateral acceleration aligns with the trend reported in section 3.1. Inferring from the baseline condition (intercept – representing the lateral acceleration of 2 m/s^2, L1 automation, and mean standardised curve radius), the odds of 2 m/s^2 being rate as Too Slow (vs Optimal) were 13.28 times higher (OR = 13.28, CI: 6.57–27.51), while the odds of it being rated as Too Fast (vs Optimal) were significantly low (OR = 0.00, CI: 0.00–0.01).

Compared to 2 m/s^2, the odds of Too Slow rating (vs Optimal) at 3 m/s^2 (OR = 0.03, CI: 0.02–0.06) and 4 m/s^2 (OR = 0.00, CI: 0.00–0.00) were much lower. In terms of Too Fast ratings (vs Optimal), 3 m/s^2 was associated with 4.65 times higher odds compared to 2 m/s^2. The odds of Too Fast ratings (vs Optimal) at 4 m/s^2 were significantly higher than at 2 m/s^2 (OR = 117.61, CI: 48.56–294.08).

The effect of automation levels.

At a lateral acceleration of 2 m/s^2, Level 4 driving (compared to L1 automation) significantly reduced the odds of Too Slow ratings

Table 3
Posterior estimation results based on multilevel Bayesian logistic model (*Rating ∼ ay + Level + Curve + ay:Level + ay:Curve + (1 | participant)*).

Parameter	Est.	SE	Est (OR)	CI.low	CI.high	Est.	SE	Est (OR)	CI.low	CI.high
Fixed effects										
Too slow (vs optimal)						*Too fast (vs optimal)*				
(Intercept)	2.59	0.37	13.28	6.57	27.51	−5.37	0.62	0.00	0.00	0.01
ay (reference: ay[2])										
ay[3]	−3.38	0.28	**0.03**	**0.02**	**0.06**	1.54	0.43	**4.65**	**2.03**	**10.97**
ay[4]	−6.21	0.40	**0.00**	**0.00**	**0.00**	4.77	0.45	**117.61**	**48.56**	**294.08**
Level (reference: Level[1])										
Level[4]	−0.62	0.25	**0.54**	**0.33**	**0.87**	−0.30	0.43	0.74	0.31	1.68
Curve radius	−2.82	0.31	**0.06**	**0.03**	**0.11**	1.41	0.52	**4.12**	**1.48**	**11.61**
ay × Level (reference: ay[2] × Level[1]) [a]										
ay[3]:Level[4]	−0.80	0.35	**0.45**	**0.22**	**0.89**	0.21	0.49	1.23	0.48	3.30
ay[4]:Level[4]	−0.39	0.44	0.67	0.28	1.60	1.41	0.47	**4.10**	**1.66**	**10.33**
ay × Curve radius (reference: ay[2] × Curve radii) [a]										
ay[3]:Curve radius	−0.92	0.40	**0.40**	**0.18**	**0.86**	1.23	0.60	**3.41**	**1.05**	**11.40**
ay[4]:Curve radius	0.09	0.56	1.10	0.35	3.24	1.94	0.56	**6.99**	**2.39**	**21.44**
Random effects										
sd(u_1)	1.54	0.22								
sd(u_2)	2.34	0.31								
cor(u_{12})	−0.66	0.11								

Note. Estimate and SE are in log-odds, while Est (OR) and Credible intervals (CI) are in odds ratio. Bold values indicate significance. Curve radii in the model were standardised. For random effects, sd(u1) and sd(u2) represent the standard deviation of the random intercepts for "Too Slow (vs Optimal)" outcome category and "Too fast (vs Optimal)" category across individuals (the grouping variable), respectively. The term cor(u12) represents the correlation between the two random intercepts across participants. a Interaction terms represent the change in predictor effects at ay = 3 m/s2 or ay = 4 m/s2 relative to the baseline combination of ay = 2 m/s2 and the respective reference level (e.g., Level = 1).

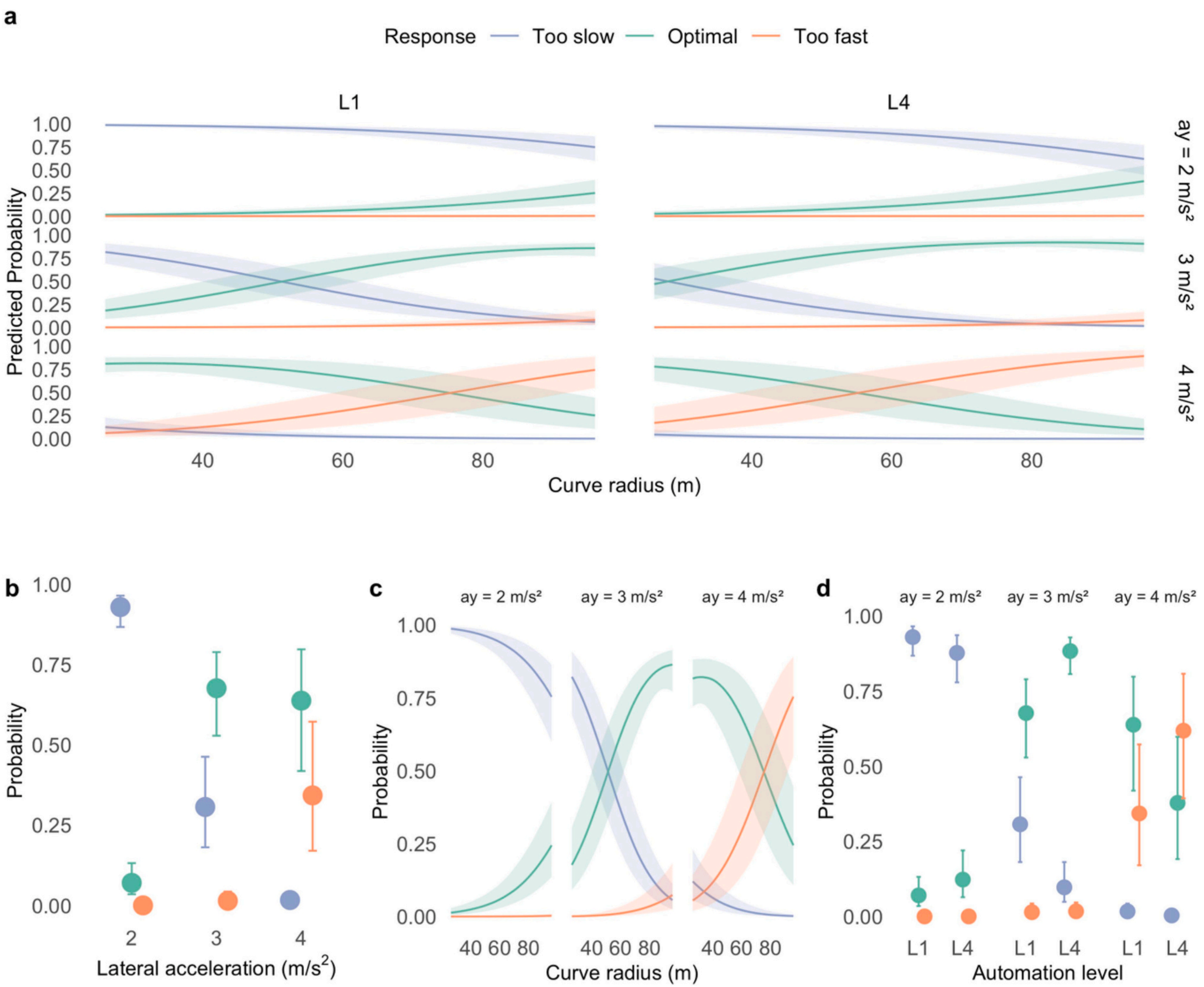

Fig. 6. Predicted probability of each rating category. (a) Predicted probabilities conditioned by lateral acceleration, automation level, and curve radius. (b) Predicted probabilities by lateral acceleration. (c) Predicted probabilities by lateral acceleration and curve radius. (d) Predicted probabilities by lateral acceleration and automation level. For plots (b)-(d), predictions were generated holding unplotted categorical predictors at their reference level and continuous predictors at their mean. For all plots, the probabilities for the three rating categories sum to 1. Error bars represent 95% CI.

(vs Optimal) by 46% (OR = 0.54, CI: 0.33–0.87). The main effect of automated level on the odds of Too Fast rating (vs Optimal) was not statistically significant (OR = 0.74, CI: 0.31–1.68).

Regarding the interaction effect, L4 automation (vs L1 automation) further reduced the odds of Too Slow ratings (vs Optimal) by 55% for 3 m/s^2, compared to its effect at 2 m/s^2 (OR = 0.45, CI: 0.22–0.89). This indicates the reduction in Too Slow odds due to L4 (vs L1) was more pronounced at 3 m/s^2 than it was at 2 m/s^2. The odds for L4 automation (vs L1) on Too Fast ratings (vs Optimal) were 4.1 times higher at 4 m/s^2 compared to its effect at 2 m/s^2 (OR = 4.10, CI: 1.66–10.33).

The effect of curve radius.

At a lateral acceleration of 2 m/s^2, a one standard deviation increase in curve radius (i.e., gentler curves) significantly reduced the odds of Too Slow ratings (vs Optimal) by 94% (OR = 0.06, CI: 0.03–0.11). In contrast, it increased the odds of Too Fast ratings (vs Optimal) 4.12 times (OR = 4.12, CI: 1.48–11.61).

Regarding interaction effects, at a lateral acceleration of 3 m/s^2, a unit increase in curve radius further reduced the odds of Too Slow (vs Optimal) ratings by 60% (OR = 0.40, CI: 0.18–0.86), compared to its effect at 2 m/s^2. Larger curve radius increased the odds of being rated as Too Fast (vs Optimal) for 3 m/s^2 (OR = 3.41, CI: 1.05–11.40) and 4 m/s^2 (OR = 6.99, CI: 2.39–21.44), compared to its effect at 2 m/s^2.

Random effects.

Similarly to the results reported in section 3.1, the random effects estimates indicate substantial between-participant variability in rating tendencies. The strong negative correlation (−0.66) between these random intercepts suggests that participants who are more likely to rate a condition Too slow (vs Optimal) are substantially less likely to rate it Too fast (vs Optimal), and vice-versa, independent of the fixed predictors.

4. Discussion

To inform the design of automated vehicles (AVs) capable of navigating varied road geometries, understanding optimal lateral acceleration settings during curve negotiation is essential. This study investigated user perception and evaluation of three specific lateral acceleration levels: 2, 3, and 4 m/s^2. These lateral accelerations, suggested by previous research to be the threshold for user comfort (Bosetti et al., 2013; J. Eriksson & Svensson, 2015; ISO 11270, 2014), were assessed across a range of conditions, including automation levels, curve radii, NDRA engagement under L4 driving, and order of rides.

4.1. Optimal lateral accelerations

Overall, the results indicate that there is *no* single universally “optimal” lateral acceleration for AVs in curve driving. Instead, passenger preference is contextual, affected by curve radii, engagement in NDRAs, their experience with automated driving, and the level of automation. However, some clear patterns have emerged which can provide design guidelines for lateral acceleration settings.

The lateral acceleration of ***2 m/s^2*** was frequently perceived as “too slow”, especially under the baseline conditions (no NDRA engagement under L4 driving, or under L1 driving). In these conditions, participants were able to look around and be aware of changes in the surrounding environment. This suggests that while safety was ensured, 2 m/s^2 may not meet passenger expectations for efficiency. In comparison, ***4 m/s^2*** consistently tended to be rated as “too fast”, as its odds of being rated as “too fast” were dramatically higher than 2 m/s^2 in both L1 and L4 (no NDRA engagement) conditions. It indicates that 4 m/s^2 appears to frequently exceed users' comfort or preference thresholds. The lateral acceleration of ***3 m/s^2*** often emerged as a more balanced option, although its ratings were dependent on other factors. Compared to 2 m/s^2, it significantly reduced the occurrence of “too slow” ratings. While it increased “too fast” ratings compared to 2 m/s^2, these increases were mostly less extreme than those seen with 4 m/s^2. Compared with literature, even the “too slow” 2 m/s^2 is much higher than the threshold of 0.98 m/s^2 proposed by de Winkel et al. (2023), which focused on participants' discomfort with discrete motion pulses when closing their eyes. It is possible that our participants mixed their evaluation for both lateral forces and vehicle speed, while the study by de Winkel focused entirely on short-period forces. Moreover, vehicle jerk might have played a role in such evaluation. Regardless, our results highlight the importance of considering vehicles' driving styles as continuous and dynamic, changing in response to external environments, to provide a full image of the scenario for participants to evaluate.

Regarding the effect of *curve radii*, larger curve radii (gentler curves) generally led to lower accelerations receiving fewer “too slow” ratings and more “too fast” ratings. Specifically, larger, gentler curves decreased “too slow” ratings and increased “too fast” ratings for 2 m/s^2, compared to smaller, sharper curves. This applied to both L1 and L4 driving. The effect on “too fast” ratings was amplified at higher accelerations, suggesting that on wider, gentler curves, passengers became more sensitive to higher lateral accelerations, perceiving them as “too fast”. This result suggests a preference for lower lateral accelerations (or lower cornering speed) on wider curves, and a preference for higher lateral accelerations (or higher cornering speed) on tighter, sharper curves. As Fig. 2 shows, at the same lateral acceleration level, the vehicle speed was higher for larger curves (e.g., Curve 2, $r = 74$ m) than for smaller and sharper curves (e.g., Curve 1, $r = 30$ m). The duration of the lateral acceleration and forces acting on the participant's body might play a role. Regardless, our results provide more detailed insights for this underexplored area about curve negotiation, where a limited number of previous studies have only inspected general trends without investigation into detailed curve radii (Hajiseyedjavadi et al., 2022; Peng et al., 2022).

Regarding *NDRA engagement in L4 driving*, visual tasks engagement led to fewer “too slow” ratings for low lateral acceleration (2 m/s^2). It suggests shifting attention from driving to NDRAs could make slow curve driving more comfortable. The significant interaction effect between NDRA engagement and curve radius indicates that engaging in visual NDRAs attenuated the effect of curve radius on both Too Slow and Too Fast ratings. Specifically, as curve radius increased, the reduction in Too Slow ratings was less pronounced under NDRA engagement compared to the No NDRA condition. This is likely because participants engaged in NDRAs were less aware of the higher vehicle speed associated with maintaining the same lateral acceleration on gentler curves. Taken together, the effect of NDRA engagement can be explained by the Multiple Resource Theory (Wickens, 2002), which suggests that human information processing relies on multiple and distinct pools of resources (e.g., visual, auditory). When a participant's visual and cognitive resources were occupied by the NDRA, their capability to process the external environment and the vehicle's motion became limited. In this case, a slower cornering speed would provide the passenger with sufficient time to perceive and evaluate the vehicle's manoeuvres. Conversely, NDRA engagement made higher lateral accelerations more likely to be rated as “too fast”. It indicates that engaging in NDRAs appears to reduce passenger comfort for higher lateral forces, possibly due to their reduced ability to monitor the vehicle's behaviour. In comparison, NDRA engagement was found not to affect preferences for longitudinal deceleration in a previous study (Carlowitz, Rossner, et al., 2024). This contrast highlights the unique influence of lateral motion on passenger perception and evaluation, and it indicates that the effect of different kinematic factors of driving styles on user evaluation warrants more in-depth research.

Regarding the effect of *automation levels*, the high automation level (L4) made low lateral acceleration (2 m/s^2) receive less “too slow” ratings compared to the lower automation level (L1). This suggests that passengers might perceive the lower lateral acceleration as more appropriate for a highly automated vehicle than for an assisted automated driving system, where they are still fundamentally the driver. L4 automation also reduced “too slow” ratings for 3 m/s^2, while it increased “too fast” ratings for 4 m/s^2. This indicates lower tolerance for high lateral acceleration in L4 driving, where the user is not in control. Taken together, our results can be explained by the user's role in different levels of automation, and their passive role in L4 driving could have contributed to the preference for lower lateral accelerations. While the experiment was not intended to induce motion sickness, the Sensory Conflict Theory suggests that discomfort arises when the brain receives conflicting signals from visual, vestibular, and proprioceptive systems (Reason, 1978). Participants' role as passengers in L4 driving limited their ability to anticipate the vehicle's upcoming movement and forces. Consequently, participants tended to prefer slower lateral acceleration to allow these multiple sensory cues to match more effectively. Moreover, the finding aligns with the general preference for defensive driving styles in highly automated driving suggested by previous research (Basu et al., 2017; Ekman et al., 2019; Peng et al., 2022).

Regarding the effect of *ride order* in L4 driving, compared with the first ride, the second ride increased “too slow” ratings for 2 m/s^2. The tendency to rate higher lateral accelerations (3 m/s^2 and 4 m/s^2) as too slow during the second ride was attenuated. While users' preferences for vehicle kinematics are likely to evolve with repeated exposure or long-term experience with highly automated driving (Metz et al., 2021), the extent to which the preferences can change still needs further investigation. Moreover, it was also possible that participants' ability to compare and evaluate across scenarios improved with practice.

4.2. Limitations and future work

While the present study offers valuable insights, certain limitations should be acknowledged, which in turn suggest avenues for future research.

First, only three discrete levels of lateral acceleration were tested in the experiment. While the choice was based on previous research and practical considerations, such as preventing participant fatigue by limiting the number of trials, it remains a limitation that intermediate values were not explored. Future research could use more granular increments to further refine the optimal zone identified in the present study and subjective preferences across different contexts.

Second, the effect of seating location within the cabin was not considered in this study, as it focused on users in the driver's seat, who could act as both drivers (in L1 driving) and passengers (in L4 driving). Passengers in other locations, particularly the backseat, are anticipated to experience greater lateral motion, due to differences in physical and ergonomic seat design. For example, front seat cushions and headrests could mitigate lateral movements more effectively than those in the back (Beard & Griffin, 2013; Reed, 2000). Moreover, backseat passengers typically have a more restricted view of the external road environment, which is considered a factor affecting comfort (Rolnick & Bles, 1989). Although future AVs may allow users to occupy the driver's seat (Koppel & Lopez-Valdes, 2019), current on-road AV models (e.g., Waymo) predominantly place users in the backseat. Therefore, the effect of seating location on user evaluations of lateral acceleration needs further investigation.

Third, the workload associated with NDRAs (the SuRT task) was consistent across conditions in this study, while the visual workload and difficulty level of the SuRT task can be varied (Guo et al., 2020). Our results suggest that NDRA engagement influenced passengers' preference for lateral acceleration, and the observed preference for lower lateral acceleration is potentially linked to perception of reduced control or situation awareness. However, research in motion sickness indicates that demanding NDRAs can reduce motion sickness, presumably by directing attention away from vehicle movements towards external stimuli (Bos, 2015). It is also important to note that the NDRA implemented in this study was a visual-manual task, requiring active, screen-based engagement. Consequently, the findings related to NDRA engagement should be interpreted as specific to this type of task, and may not generalise to NDRAs of a different modality (e.g., auditory tasks). To comprehensively understand the effect of varying cognitive workloads and NDRA modality on passengers' preferences for lateral acceleration, further research is needed.

Fourth, our primary aim was to establish the optimal range of lateral acceleration and how contextual factors shift the perception of that specific variable. Consequently, we intentionally limited our interaction effects to those involving lateral acceleration. However, future research with larger sample sizes could explore three-way interactions to further refine the understanding of comfortable lateral acceleration.

Lastly, our finding that participants preferred higher lateral acceleration for sharper curves might be explained by the study design and the scenario. Specifically, maintaining target lateral accelerations on sharper curves necessitated relatively lower vehicle speeds, while the test track was a racing track type. This speed constraint was also to ensure the safety of the ride due to the winter conditions during the experiment. Future studies could explore this relationship by varying road types (e.g., urban vs motorway) and adjusting vehicle speed for different road types, to better disentangle the relationship between preferred lateral acceleration, vehicle speed, and curve radii. Moreover, the presence of a safety driver and the absence of surrounding traffic in the test track necessitate more investigations.

5. Conclusion

This study investigated how passengers perceive and evaluate lateral acceleration in automated vehicles negotiating curves, across varying curve radii, NDRA engagement, and automation levels. Within the tested lateral acceleration levels of 2, 3 and 4 m/s^2, preferred lateral acceleration depended on contextual factors including curve geometry, NDRA engagement, and automation level. Within the tested range, while 3 m/s^2 frequently emerged as the most balanced of the tested acceleration levels, its acceptability depended on curve geometry, NDRA engagement, and automation level. Ratings of both 2 m/s^2 and 4 m/s^2 generally showed greater deviations from the optimal category. These findings have implications for the design of AV lateral control systems. First, regarding curve geometry, passengers preferred lower lateral accelerations on gentler curves with larger curve radii, where the vehicle travelled at higher speeds, while they were more tolerant of higher accelerations on sharper curves This pattern reflects the relationship between speed, curve radius, and lateral acceleration, suggesting that AV controllers should adapt lateral acceleration targets to curve geometry rather than applying a fixed value across all curves. Second, regarding NDRA engagement, engaging in visual tasks attenuated the moderating effect of curve radius on motion perception, suggesting that passengers doing NDRAs are less sensitive to the higher speed associated with gentler curves. AV systems that can detect passenger engagement in NDRAs could therefore adopt less conservative lateral acceleration levels on gentler curves when passengers are engaged in NDRAs, although this warrants further investigation into the balance between safety, comfort and trust. Third, regarding automation level, passengers under L4 automation were more accepting of lower accelerations and less tolerant of higher accelerations compared to L1, consistent with their passive role and reduced ability to anticipate vehicle manoeuvres. This suggests that lateral acceleration profiles should be calibrated differently depending on the active automation level.

Taken together, these findings suggest a context-aware system that dynamically adjust lateral acceleration based on road geometry, passenger activity, and automation level may be better able to maintain consistent passenger comfort across a wide range of real-world driving conditions.

CRediT authorship contribution statement

Chen Peng: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization. **Stefanie Carlowitz:** Writing – review & editing, Writing – original draft, Investigation. **Ruth Madigan:** Writing – review & editing. **Claus Marberger:** Writing – review & editing, Resources, Investigation, Conceptualization. **Michael Schulz:** Writing – review & editing, Resources, Project administration, Investigation, Data curation, Conceptualization. **Sabine Osswalt:** Software, Investigation. **Andreas Schultz:** Software, Investigation. **Natasha Merat:** Writing – review & editing, Supervision, Project administration.

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.

Acknowledgement

This study was conducted within the RUMBA project (www.projekt-rumba.de, accessed on 9th October 2024), which is funded by the German Federal Ministry of Economic Affairs and Climate Action (funding number: 19A20007K).

This study was also funded by the Hi-Drive project, as it provided support for the time of CP (the first author), RM and NM. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101006664. The sole responsibility of this publication lies with the authors. Neither the European Commission nor CINEA – in its capacity of Granting Authority – can be made responsible for any use that may be made of the information this document contains.

Appendix 1

As the data distribution shows (Fig. 7), while the trend of ratings for lateral acceleration levels remains similar, the engagement in NDRAs and different curves have a certain impact on subjective ratings. Four mixed-effects models were fitted: Model 1 only included lateral acceleration; Model 2 included lateral acceleration, NDRA engagement and ride order; Model 3 included curves in addition to predictors in Model 2; and Model 4 included further two-way interaction terms. Based on the comparison of model performance

(Table 4), we report details for Model 4).

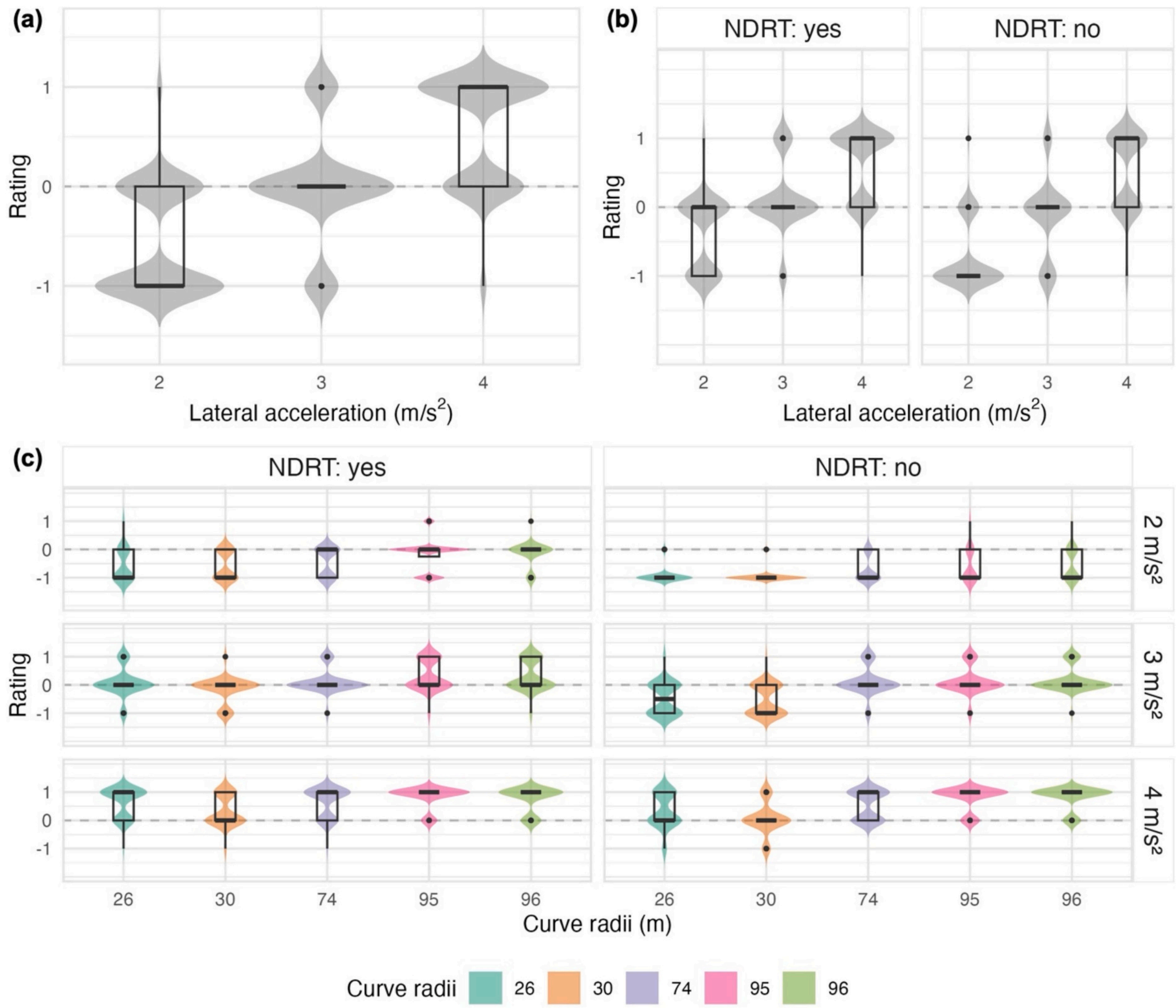

Fig. 7. Data distribution of ratings for lateral acceleration levels in L4 automated driving, in terms of combining all conditions (a), considering NDRA engagement (b), and considering both NDRA engagement and curve radii (c).

Table 4
Comparison of model performance.

Model	ELPD diff	SE diff
Model 4: *Rating ~ ay + NDRA + Curve + order + ay:NDRA + ay:Curve + ay:order + NDRA:Curve + NDRA:order + Curve:order + (1 | participant)*	0	0
Model 3: Rating ~ ay + NDRA + Curve + order + ay:NDRA + ay:Curve + ay:order + (1 | participant)	−8.1	3.7
Model 2: Rating ~ ay + NDRA + ay:NDRA + order + ay:order + (1 | participant)	−204.2	16.7
Model 1: Rating ~ ay + (1 | participant)	−277.5	19.4

Note. A higher value of expected log predictive density (ELPD) means better predictive accuracy. SE represents the standard error of the difference in ELPD between each pair of models.

Appendix 2

To compare the evaluation of lateral acceleration between L1 and L4 automation, the condition where participants engaged in NDRA in L4 automated driving was excluded. Fig. 8 shows the distribution of the data. Four models were tested (Table 5), and Model 3, including lateral acceleration, automation level, and curve radii, was selected and reported in detail. The factor of the ride order did not improve model fit and thus was not included.

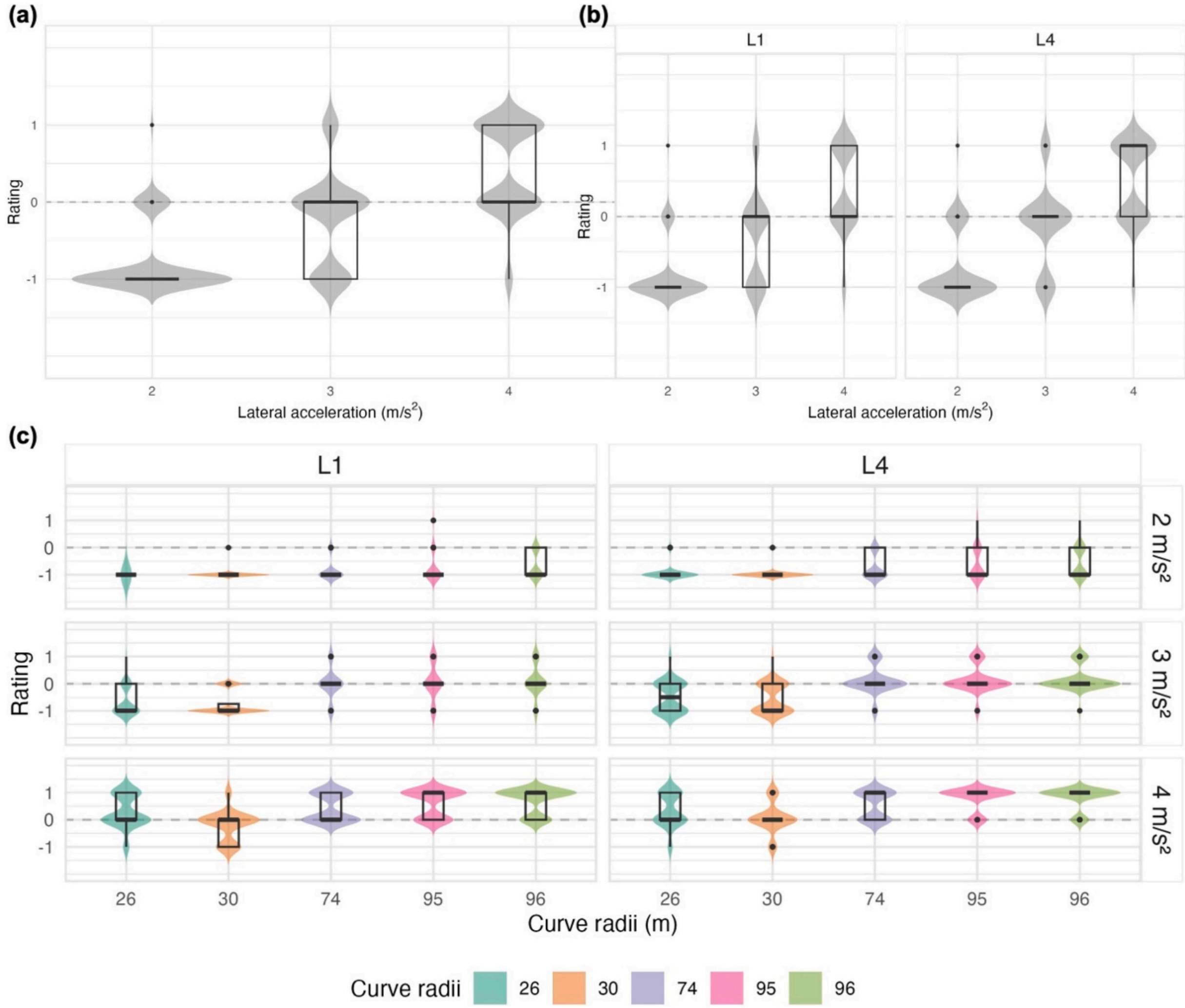

Fig. 8. Data distribution of ratings for lateral acceleration levels in different levels of automated driving, in terms of combining all conditions (a), considering automation level (b), and considering both automation level and curve radii (c).

Table 5
Comparison of model performance.

Model	ELPD diff	SE diff
Model 3: Rating ∼ ay + level + curve + ay:level + ay:curve + (1 | participant)	0	0
Model 4: Rating ∼ ay + level + curve + ay:level + ay:curve + level:curve + (1 | participant)	−1.4	1.3
Model 2: Rating ∼ ay + level + ay:level + (1 | participant)	−319.3	19.4
Model 1: Rating ∼ ay + (1 | participant)	−334.3	19.9

Note. A higher value of expected log predictive density (ELPD) means better predictive accuracy. SE represents the standard error of the difference in ELPD between each pair of models.

Data availability

The data that has been used is confidential.

References

Agresti, A. (2003). *Categorical data analysis, Wiley-Interscience* (2nd ed.).

Bae, I., Moon, J., Jhung, J., Suk, H., Kim, T., Park, H., Cha, J., Kim, J., Kim, D., & Kim, S. (2022). Self-driving like a human driver instead of a robocar: Personalized comfortable driving experience for autonomous vehicles (arXiv:2001.03908). *arXiv*. https://doi.org/10.48550/arXiv.2001.03908

Bae, I., Moon, J., & Seo, J. (2019). Toward a comfortable driving experience for a self-driving shuttle bus. *Electronics (Switzerland), 8*(9), 1–13. https://doi.org/10.3390/electronics8090943

Basu, C., Yang, Q., Hungerman, D., Singhal, M., & Dragan, A. D. (2017). *Do you want your autonomous Car to drive like you?*. ACM/IEEE International Conference on Human-Robot Interaction, Part F1271, (pp. 417–425). https://doi.org/10.1145/2909824.3020250.

Bayes, T. (1763). LII. An essay towards solving a problem in the doctrine of chances. By the late rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society of London, 53*, 370–418. https://doi.org/10.1098/rstl.1763.0053

Beard, G. F., & Griffin, M. J. (2013). Discomfort during lateral acceleration: Influence of seat cushion and backrest. *Applied Ergonomics, 44*(4), 588–594. https://doi.org/10.1016/j.apergo.2012.11.009

Beggiato, M., Hartwich, F., Schleinitz, K., Krems, J. F., Othersen, I., & Petermann-Stock, I. (2015). *What would drivers like to know during automated driving? Information needs at different levels of automation*. . 7thConference on Driver Assistance,. 49. 0 (pp. 1–7). https://doi.org/10.13140/RG.2.1.2462.6007.

Bos, J. (2015). Less sickness with more motion and/or mental distraction. *Journal of Vestibular Research : Equilibrium & Orientation, 25*, 23–33. https://doi.org/10.3233/VES-150541

Bosetti, P., Lio, M., & Saroldi, A. (2013). On the human control of vehicles: An experimental study of acceleration. *European Transport Research Review, 6*. https://doi.org/10.1007/s12544-013-0120-2

Brietzke, A., Pham Xuan, R., Dettmann, A., & Bullinger, A. C. (2021). Influence of dynamic stimulation, visual perception and individual susceptibility to car sickness during controlled stop-and-go driving. *Forschung im Ingenieurwesen, 85*(2), 517–526. https://doi.org/10.1007/s10010-021-00441-6

Bürkner, P.-C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*, 1–28. https://doi.org/10.18637/jss.v080.i01

Carlowitz, S., Madigan, R., Goodridge, C. M., Hilz, J., Marberger, C., Alt, P., … Merat, N. (2024). Balancing comfort deceleration rate and environmental impact at crosswalks and intersections in automated driving. *SSRN scholarly paper, 5063808*. https://doi.org/10.2139/ssrn.5063808

Carlowitz, S., Rossner, P., Madigan, R., Bieg, H.-J., Marberger, C., Alt, P., … Merat, N. (2024). User evaluation of comfortable deceleration profiles for highly automated driving: Findings from a test track study. *Transportation Research Part F: Traffic Psychology and Behaviour, 105*, 206–221. https://doi.org/10.1016/j.trf.2024.05.025

Chen, C., Zhang, G., Liu, X. C., Ci, Y., Huang, H., Ma, J., … Guan, H. (2016). Driver injury severity outcome analysis in rural interstate highway crashes: A two-level Bayesian logistic regression interpretation. *Accident Analysis & Prevention, 97*, 69–78. https://doi.org/10.1016/j.aap.2016.07.031

Delbiso, T. D., Rodriguez-Llanes, J. M., Altare, C., Masquelier, B., & Guha-Sapir, D. (2016). Health at the borders: Bayesian multilevel analysis of women's malnutrition determinants in Ethiopia. *Global Health Action, 9*(1), Article 30204. https://doi.org/10.3402/gha.v9.30204

Eboli, L., Mazzulla, G., & Pungillo, G. (2016). *Measuring bus comfort levels by using acceleration instantaneous values. Transportation research Procedia, efficient, safe and intelligent transport. Selected papers from the XII conference on transport engineering, Valencia (Spain). 18* pp. 27–34). https://doi.org/10.1016/j.trpro.2016.12.004

Ekman, F., Johansson, M., Bligård, L. O., Karlsson, M. A., & 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. https://doi.org/10.1016/j.trf.2019.07.026

Elbanhawi, M., Simic, M., & Jazar, R. (2015). In the passenger seat: Investigating ride comfort measures in autonomous cars. *IEEE Intelligent Transportation Systems Magazine, 7*(3), 4–17. https://doi.org/10.1109/MITS.2015.2405571

Eriksson, J., & Svensson, L. (2015). Tuning for ride quality in autonomous vehicle application to linear quadratic path planning algorithm. . In *Master Thesis,*Uppsala Universitet http://www.teknat.uu.se/student.

Eriksson, P., & Friberg, O. (2000). Ride comfort optimization of a city bus. *Structural and Multidisciplinary Optimization, 20*(1), 67–75. https://doi.org/10.1007/s001580050137

Goldstein, H. (2011). *Multilevel Statistical Models, John Wiley & Sons*.

Gudmestad, A., & Metzger, T. A. (2025). Mixed-effects modeling with a multinomial dependent variable. *Language Learning, 75*(1), 212–249. https://doi.org/10.1111/lang.12667

Guo, H., Boyle, L. N., Jenness, J. W., & Lee, J. D. (2020). Tactile detection response task: Metrics for assessing drivers' cognitive workload. *Transportation Research Part F: Traffic Psychology and Behaviour, 70*, 98–108. https://doi.org/10.1016/j.trf.2019.12.003

Hajiseyedjavadi, F., Boer, E. R., Romano, R., Paschalidis, E., Wei, C., Solernou, A., … Merat, N. (2022). Effect of environmental factors and individual differences on subjective evaluation of human-like and conventional automated vehicle controllers. *Transportation Research Part F: Traffic Psychology and Behaviour, 90*, 1–14. https://doi.org/10.1016/j.trf.2022.07.018

Hartwich, F., Beggiato, M., & Krems, J. F. (2018). Driving comfort, enjoyment and acceptance of automated driving–effects of drivers' age and driving style familiarity. *Ergonomics, 61*(8), 1017–1032. https://doi.org/10.1080/00140139.2018.1441448

Hedeker, D. (2003). A mixed-effects multinomial logistic regression model. *Statistics in Medicine, 22*(9), 1433–1446. https://doi.org/10.1002/sim.1522

Heiss, A. (2023). *The ultimate practical guide to multilevel multinomial conjoint analysis with R. Andrew Heiss*. https://doi.org/10.59350/2mz75-rrc46

ISO 11270. (2014). ISO *11270:2014, Intelligent transport systems—Lane keeping assistance systems (LKAS)—Performance requirements and test procedures*. https://www.iso.org/obp/ui/#iso:std:iso:11270:ed-1:v1:en.

ISO 15622. (2018). Intelligent transport systems—Adaptive cruise control systems—Performance requirements and test procedures. https://www.iso.org/standard/71515.html.

ISO 17488:2016, ISO. https://www.iso.org/standard/59887.html, (2016).

Kato, K., & Kitazaki, S. (2006). *A study for understanding carsickness based on the sensory conflict theory* (SAE technical paper Nos 2006-01–0096). SAE Technical Paper. Doi: https://doi.org/10.4271/2006-01-0096.

Koppel, S., Octavio, J., Bohman, J., Logan, K., Raphael, D., Quintana, W., … Lopez-Valdes, F. (2019). Seating configuration and position preferences in fully automated vehicles. *Traffic Injury Prevention, 20*(sup2), S103–S109. https://doi.org/10.1080/15389588.2019.1625336

Kyriakidis, M., de Winter, J. C. F., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., … Happee, R. (2019). A human factors perspective on automated driving. *Theoretical Issues in Ergonomics Science, 20*(3), 223–249. https://doi.org/10.1080/1463922X.2017.1293187

Madigan, R., Louw, T., & Merat, N. (2018). The effect of varying levels of vehicle automation on drivers' lane changing behaviour. *PLoS One, 13*(2), Article e0192190. https://doi.org/10.1371/journal.pone.0192190

Martin, D., & Litwhiler, D. (2008). *An investigation of acceleration and jerk profiles of public transportation vehicles*. . *ASEE Annual Conference and Exposition, Conference Proceedings*, (1984).

Metz, B., Wörle, J., Hanig, M., Schmitt, M., Lutz, A., & Neukum, A. (2021). Repeated usage of a motorway automated driving function: Automation level and behavioural adaption. *Transportation Research Part F: Traffic Psychology and Behaviour, 81*, 82–100. https://doi.org/10.1016/j.trf.2021.05.017

Metzulat, M., Metz, B., Landau, A., Neukum, A., & Kunde, W. (2024). Does the visual input matter? Influence of non-driving related tasks on car sickness in an open road setting. *Transportation Research Part F: Traffic Psychology and Behaviour, 104*, 234–248. https://doi.org/10.1016/j.trf.2024.06.002

Miller, D., Sun, A., & Ju, W. (2014). *Situation awareness with different levels of automation*. . *2014 IEEE international conference on systems, man, and cybernetics (SMC)*, (pp. 688–693). https://doi.org/10.1109/SMC.2014.6973989.

Mühl, K., Strauch, C., Grabmaier, C., Reithinger, S., Huckauf, A., & Baumann, M. (2019). Get ready for being chauffeured: Passenger's preferences and trust while being driven by human and automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. https://doi.org/10.1177/0018720819872893, 001872081987289.

Ossig, J., Cramer, S., Eckl, A., & Bengler, K. (2023). Tactical decisions for lane changes or lane following: Assessment of automated driving styles under real-world conditions. *IEEE Transactions on Intelligent Vehicles, 8*(1), 502–511. https://doi.org/10.1109/TIV.2022.3161570

Paries, M., Bougeard, S., & Vigneau, E. (2022). Multivariate analysis of just-about-right data with optimal scaling approach. *Food Quality and Preference, 102*, Article 104681. https://doi.org/10.1016/j.foodqual.2022.104681

Peng, C. (2024). Understanding and improving user comfort in automated driving [Phd, University of Leeds]. https://etheses.whiterose.ac.uk/34995/

Peng, C., Horn, S., Madigan, R., Marberger, C., Lee, J. D., Krems, J., … Merat, N. (2024). Conceptualising user comfort in automated driving: Findings from an expert group workshop. *Transportation Research Interdisciplinary Perspectives, 24*, Article 101070. https://doi.org/10.1016/j.trip.2024.101070

Peng, C., Merat, N., Romano, R., Hajiseyedjavadi, F., Paschalidis, E., Wei, C., Radhakrishnan, V., Solernou, A., Forster, D., & Boer, E. (2022). Drivers' evaluation of different automated driving styles: Is it both comfortable and natural? *Human Factors.*. https://doi.org/10.1177/00187208221113448

Peng, C., Wei, C., Solernou, A., Hagenzieker, M., & Merat, N. (2025). User comfort and naturalness of automated driving: The effect of vehicle kinematic and proxemic factors on subjective response. *Applied Ergonomics, 122*, Article 104397. https://doi.org/10.1016/j.apergo.2024.104397

Price, M., Venkatraman, V., Gibson, M., Lee, J., & Mutlu, B. (2016). Psychophysics of trust in vehicle control algorithms. *SAE Technical Paper*. https://doi.org/10.4271/2016-01-0144

Reason, J. T. (1978). Motion sickness adaptation: A neural mismatch model. *Journal of the Royal Society of Medicine, 71*(11), 819–829. https://doi.org/10.1177/014107687807101109

Reed, M. P. (2000). *Survey of auto seat design recommendations for improved comfort*.

Rolnick, A., & Bles, W. (1989). Performance and well-being under tilting conditions: The effects of visual reference and artificial horizon. *Aviation, Space, and Environmental Medicine, 60*(8), 779–785.

Rothman, L., & Parker, M. J. (2009). Structure and use of just-about-right scales. In L. Rothman, & M. J. Parker (Eds.), *Just-about-right (JAR) scales: Design, usage, benefits, and risks, ASTM International* (p. 0). https://doi.org/10.1520/MNL11481M

SAE. (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*.

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., … Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers, 1*(1), 1–26. https://doi.org/10.1038/s43586-020-00001-2

Stayton, E., & Stilgoe, J. (2020). It's time to rethink levels of automation for self-driving vehicles [opinion]. *IEEE Technology and Society Magazine, 39*(3), 13–19. https://doi.org/10.1109/MTS.2020.3012315

Tinga, A. M., van Zeumeren, I. M., Christoph, M., van Grondelle, E., Cleij, D., Aldea, A., & van Nes, N. (2023). Development and evaluation of a human machine interface to support mode awareness in different automated driving modes. *Transportation Research Part F: Traffic Psychology and Behaviour, 92*, 238–254. https://doi.org/10.1016/j.trf.2022.10.023

Turner, M., & Griffin, M. J. (1999). Motion sickness in public road transport: The effect of driver, route and vehicle. *Ergonomics, 42*(12), 1646–1664. https://doi.org/10.1080/001401399184730

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing, 27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Wang, C., Zhao, X., Fu, R., & Li, Z. (2020). Research on the comfort of vehicle passengers considering the vehicle motion state and passenger physiological characteristics: Improving the passenger comfort of autonomous vehicles. *International Journal of Environmental Research and Public Health, 17*(18), Article 18. https://doi.org/10.3390/ijerph17186821

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159–177. https://doi.org/10.1080/14639220210123806

de Winkel, K., Irmak, T., Happee, R., & Shyrokau, B. (2023). Standards for passenger comfort in automated vehicles: Acceleration and jerk. *Applied Ergonomics, 106*, Article 103881. https://doi.org/10.1016/j.apergo.2022.103881

Young, R. A., Hsieh, L., & Seaman, S. (2013). *The tactile detection response task: Preliminary validation for measuring the attentional effects of cognitive load.* . Driving assessment conference,. 7. (2013), article 2013. Doi: https://doi.org/10.17077/drivingassessment.1469.

Yusof, N. M., Karjanto, J., Terken, J., Delbressine, F., Hassan, M. Z., & Rauterberg, M. (2016). *The exploration of autonomous vehicle driving styles: Preferred longitudinal, lateral, and vertical accelerations*. . AutomotiveUI 2016 - 8th international conference on automotive user interfaces and interactive vehicular applications, proceedings, (pp. 245–252). https://doi.org/10.1145/3003715.3005455.