Human-Computer Interaction & Emerging Technologies, Vol. 233, 2026, 362–372
https://doi.org/10.54941/ahfe1007536

From Ambiguous Intentions to Reassuring Yielding: A Chinese Cultural Tradition "Li" Based Interaction Etiquette Model for Vehicle eHMI Design

Nan Wang and Chunxi Huang

School of Design, Hunan University, Changsha 410082, China

ABSTRACT

Rooted in the Chinese cultural concept of Li (etiquette), this study addresses the challenge of ambiguous intention communication in vehicle–pedestrian interactions in the AI era. While existing eHMIs emphasize functional clarity, they often fail to provide the emotional reassurance necessary for trust formation in mixed traffic environments. To bridge this gap, we adopted a qualitative, bottom-up approach, collecting over 50,000 raw entries from major Chinese social platforms and analyzing 1,229 valid samples using Grounded Theory. This analysis resulted in an Interaction Etiquette Model structured around three human factors dimensions—Vitality, reflecting perceived attentiveness and temporal rhythm; Norm, representing predictability and rule compliance; and Modesty, capturing non-intrusiveness and affective restraint—which together explain how culturally embedded expectations shape pedestrian trust calibration. Building on this model, we propose an AI–Aesthetics dual-drive framework in which AI enables accurate perception and context-aware decision-making, while aesthetic design translates these capabilities into socially legible and emotionally reassuring interaction cues. Based on this framework, we outline strategies for visual and personalized intention signaling, such as graded salience and rhythmic lighting, to reduce uncertainty. Overall, this work contributes a culturally grounded design paradigm for eHMI, extending existing approaches from functional signaling toward socially adaptive, trust-oriented interaction design.

Keywords: Vehicle-pedestrian interaction, eHMI, Interaction aesthetics, Grounded theory, Trust

INTRODUCTION

The transition from SAE Level 2 (Partial Automation) to Level 4 (High Automation) fundamentally alters the cognitive dynamics of road safety. In traditional traffic, vehicle-pedestrian interaction (VPI) relies heavily on informal social signaling—eye contact, hand gestures, and subtle vehicle kinematics—that allow pedestrians to infer driver intent and negotiate right-of-way (Rasouli and Tsotsos, 2019). The removal of the human driver creates a "communication vacuum," forcing pedestrians to rely solely on the vehicle's physical behavior and external interfaces (eHMI) to make critical crossing decisions (Mahadevan et al., 2018).

From a human factors perspective, this shift introduces significant cognitive load and uncertainty (Tabone et al., 2021). Pedestrians must construct new

Received February 5, 2026; Revised April 5, 2026; Accepted April 20, 2026; Available online July 20, 2026

mental models for "machine behavior," often leading to over-trust (assuming the vehicle sees them) or under-trust (hesitation and paralysis). While current eHMI research emphasizes signal visibility and standardized iconography, it often overlooks the social legibility of interaction—how a vehicle's behavior aligns with cultural expectations of politeness and safety (Colley and Rukzio, 2020).

Cultural Ergonomics: "Li" as a Cognitive Schema

Trust in automated systems is not merely a technical calculation but a socially situated cognitive process shaped by norms, expectations, and prior interaction experience (Hoff and Bashir, 2015). In dense, mixed-traffic environments such as Chinese cities, traffic negotiation is strongly influenced by implicit social norms emphasizing restraint, reciprocity, and situational appropriateness—principles that resonate with the Confucian concept of Li (礼).

Within a human factors framework, such norms can be understood as shared cognitive schemas that reduce interpretive ambiguity and interactional friction, consistent with research in cultural ergonomics and cross-cultural HCI (Rouchitsas and Alm, 2019). When an automated vehicle behaves in ways that violate these expectations—for example, by braking abruptly without anticipatory cues or by waiting passively without communicative feedback—pedestrians may experience cognitive dissonance, discomfort, or mistrust, despite objective safety (Stampf et al., 2022).

Accordingly, this study argues that achieving reassuring yielding requires automated vehicles to align not only with traffic rules but also with culturally grounded expectations of polite and orderly conduct. By translating such norms into interaction-relevant behavioral patterns, AVs may transform "cold" algorithmic decisions into socially legible and emotionally acceptable actions.

Research Objectives

This study aims to bridge the gap between the technical functionality of autonomous vehicles and their social acceptance in vehicle–pedestrian interaction by introducing a design perspective grounded in the Confucian concept of Li (ritual and etiquette).

This study makes three key contributions to the field:

(1) Conceptualization: It operationalizes the abstract concept of Li into three measurable human factors dimensions—Vitality, Norm, and Modesty;
(2) Modeling: It develops an Interaction Etiquette Model that explains how culturally embedded expectations influence pedestrian trust calibration;
(3) Design Framework: It proposes an AI–Aesthetics dual-drive framework that translates machine intelligence into socially legible and emotionally reassuring interaction cues.

By doing so, this study positions itself at the intersection of eHMI design for autonomous vehicles in the AI era, cultural ergonomics, and trust research in automated mobility.

THEORETICAL BACKGROUND

eHMI as a Cognitive Interface

External Human-Machine Interfaces (eHMI) serve as the primary channel for AVs to communicate intent, replacing the social cues traditionally provided by human drivers. Current research categorizes eHMI into visual (lights, displays, projections), auditory, and kinematic modalities, with visual carriers being the most direct form of communication for explicit intent signaling (Eisele et al., 2025).

Trust Calibration in Automation

Trust calibration is defined as the alignment of a user's trust in a system with the system's actual capabilities. In VPI, "calibrated trust" means pedestrians correctly identify when it is safe to cross and when it is not (Islam and She, 2025). Over-trust leads to dangerous crossings (e.g., assuming an AV will stop when it cannot), while under-trust leads to traffic inefficiency and hesitation (Hensch et al., 2022). Our theoretical premise is that Li-based aesthetics serve as "trust cues" that help pedestrians calibrate their expectations more accurately than raw kinematic data alone, mitigating the risks of over-trust in ambiguous scenarios (Mahmood and Szabolcsi, 2025). However, existing studies primarily focus on functional clarity and signal recognition, with limited attention to culturally grounded interaction expectations, which motivates the present study.

METHODOLOGY

Grounded Theory Research Process

Grounded Theory, proposed by Glaser et al. in 1967, is a qualitative methodology whose core lies in "growing theory from data." Grounded Theory requires researchers not to preset theoretical hypotheses but to systematically collect and analyze data from specific subjects, discover basic concepts, grasp internal logical relationships among concepts, and then establish explanatory theoretical frameworks. Recent studies in human-factors engineering have successfully applied Grounded Theory to model complex trust dynamics in automated vehicle interactions, demonstrating its efficacy in capturing user-centric safety needs (Sudhakaran et al., 2025). Existing Grounded Theory research often utilizes text data such as online reviews and in-depth interviews for analysis. Online reviews, as immediate public feedback in real traffic scenarios, possess characteristics such as large data volume, authentic emotions, and rich scenarios, serving as important channels for analyzing human-vehicle interaction psychology and reflecting public trust needs toward autonomous driving (Sujon and Dai, 2021).

Therefore, this study uses multi-platform online reviews as data sources, employing NVivo 15 qualitative analysis software as the research tool. After importing cleaned review data into the software, strictly following Grounded Theory coding procedures (open–axial–selective) and combining word frequency statistics and visualization tools, it completes the full process of concept extraction, category induction, model construction, and saturation testing. Data collection and theory construction followed a constant comparative and iterative process until theoretical saturation was reached, enabling a deeper exploration of the underlying mechanisms of Li culture in traffic interaction and enhancing the credibility and rigor of the study.

Online Review-Related Research and Data Sources

Online reviews refer to personal viewpoints and emotional expressions published by users on network platforms after experiencing specific events or using products, objectively reflecting users' authentic experiences and psychological appeals. In transportation research, traditional questionnaire surveys often struggle to capture subtle psychology during drivers' and pedestrians' instantaneous strategic interactions, whereas online reviews on social media can authentically record these "unstructured" interaction details (Lian et al., 2020). Currently, mainstream domestic and international research has begun utilizing online reviews to analyze public acceptance and trust toward autonomous driving technology. For instance, recent studies have leveraged social media sentiment analysis to construct adoption models for autonomous vehicles, revealing the critical role of public discourse in shaping trust (Vafaei-Zadeh et al., 2025) while subjective norms positively influence both perceived usefulness and perceived enjoyment (PEN. These reviews provide intuitive and thorough public feedback for this study. Therefore, this study selects online reviews as the primary data source for exploring driver-pedestrian interaction behaviors in the Chinese context.

Data Collection and Organization Based on Web Scraping Technology

As indicated above, existing research urgently needs insights into the impact of social norms, value identity, regional differences, and non-verbal symbols on human-vehicle trust. This study uses keywords such as "vehicles yielding to people," "yielding to pedestrians," "people yielding to vehicles," and "yielding vehicle interaction" to systematically scrape relevant user comments from four major platforms—Xiaohongshu, Weibo, Zhihu, and Autohome—via the Octopus web scraping program, exploring public interaction patterns and psychological expectations in real traffic scenarios. Given the massive volume of raw data and the presence of substantial irrelevant information, this study selects user comments from the past decade (2015–2025) for screening, scraping over 50,000 raw data entries. After cleaning by removing promotional advertisements, duplicate comments, and meaningless statements (e.g., pure emojis), 23,967 entries were eliminated. The remaining entries were imported into NVivo 15 for word frequency statistics and word cloud generation (see Figure 1). Based on word frequency statistics and word cloud distribution, Python tools were employed through comment text word counts, keyword matching, and information density screening to obtain 1,229 core data entries strongly related to the "yielding to pedestrians + human-machine interaction" theme with high information content, serving as the "core dataset" for subsequent deep coding (Sujon and Dai, 2021). Additionally, to ensure theoretical construction rigor, 20 valid comments were randomly extracted before coding as reserved data for subsequent theoretical saturation testing, while the remaining 1,209 comments were used to construct the "Interaction Etiquette" framework model.

Figure 1: Valid data word cloud.

Data Coding and Model Construction

Open Coding

First, open coding was performed on the data. Since original corpus volumes are large and unstructured, line-by-line splitting and refinement are necessary for labeling, conceptualization, and categorization of raw materials. This rigorous process of breaking down qualitative data into discrete parts to identify core themes is a standard practice in transportation safety research (Channamallu et al., 2025). By analyzing 1,209 comments line-by-line in "ID-Comment Content" format, we extracted 1,223 reference nodes. These nodes represent an initial, low-level coding structure, directly reflecting detailed elements of human-vehicle interaction. Next, categorization analysis was performed on these nodes. Referring to NVivo-generated word frequency clouds, nodes with identical semantic orientations were merged into initial concepts. For example, a group of reviews stated: "Hangzhou's yielding is well-implemented; highways are spontaneous," "Last time in Beijing, passing a small intersection without traffic lights, I thought cars would yield, but the car didn't brake at all." These reviews were uniformly labeled "regional differences," a factor increasingly recognized in cross-cultural AV interaction studies (Xin et al., 2025). "Small cars and private cars don't yield, honking horns to force pedestrians to stop and rush through themselves," "Must drive forward, desperately honking horns forcing people to dodge" were labeled "vehicle language," aligning with recent research on the "Language of Driving" and explicit acoustic signals (Kalda et al., 2022). Following this method, 21 initial categories were derived, including "regional differences," "vehicles yielding to people," "vehicle language," and "social norms."

Axial Coding

Axial coding, based on open coding summarization, analyzes causal, situational, and interactive relationships among different initial categories, inducing major categories and clarifying internal logic among concepts. This phase is critical for reconstructing data into a coherent theoretical model (Sosik-Filipiak and Osypchuk, 2023). The 21 initial categories were summarized according to "Situation Construction →Trust Decision →Dynamic Interaction →Emotional Feedback," mirroring the sequential state models proposed in recent driver-pedestrian communication research (Pipkorn et al., 2025). For example, three initial concepts formed by open coding—"institutional trust," "vehicle behavior norms," and "pedestrian behavior norms"—can be integrated into the major category "rule compliance"; "road type" and "regional differences" can be integrated into "situational adaptation." Ultimately, 8 intermediate-stage main categories were derived as key factors influencing human-vehicle interaction trust, i.e., rule compliance, boundary definition, situational adaptation, behavioral expectation, dynamic capability, rhythm coordination, etiquette yielding, and non-verbal communication.

Selective Coding

Selective coding refines and integrates previous-level coding, identifying core categories from major categories and excavating the logical relationships between major and minor categories, constructing a theoretical framework capable of clearly describing phenomena. Based on analyzing the above 8 main categories combined with traditional Chinese cultural contexts, researchers constructed a targeted Grounded Theory system: "Interaction Etiquette Model." This model summarizes human-vehicle interaction core logic into three core aesthetic dimensions: "Vitality-Norm-Modesty." These three dimensions, moderated by situational adaptation and driven by cultural psychology, jointly act on trust perception formation. These five factors mutually influence and interact, forming a "Situation–Behavior–Trust" feedback loop structure. Selective coding and main category relationship structures are shown in Table 1.

Table 1: Selective coding.

Initial Concept	Category (Node)	Main Category	Definition
Institutional trust (a1); Pedestrian behavior norms (a2); Vehicle behavior norms (a3); Vehicles yielding to people (a4); Vehicles yielding to vehicles (a5); Vehicles not yielding (a6); Road type (a7); Optimization suggestions (a10); Predictable trust (a11)	Rule compliance (361); Boundary definition (308); Behavioral expectation (44)	Norm (58.30%)	Interaction "hard constraints," emphasizing objective rules, physical boundaries, and institutional safeguards

(Continued)

Table 1: Continued.

Initial Concept	Category (Node)	Main Category	Definition
Pedestrian crossing speed (a13); Pedestrian hesitation (a14); Vehicle urging (a15); Regional differences (a8); Individual characteristics (a9); Capability trust (a12)	Rhythm coordination (43); Situational adaptation (263); Dynamic capability (4)	Vitality (25.35%)	Interaction temporal rhythm, emphasizing "no rushing, no dragging" flow and continuity
People yielding to vehicles (a16); Goodwill trust (a17); Vehicle language (a18); Body language (a19); Pedestrian feedback (a20); Li culture value (a21)	Etiquette yielding (65); Non-verbal communication (48); Emotional interaction (87)	Modesty (16.35%)	Interaction behavioral attitude, emphasizing "mutual yielding" humility and goodwill expression

Theoretical Saturation Testing

Following theory construction, a theoretical saturation test was conducted on 20 reserved samples. No new concepts, categories, or relationships emerged beyond the established three dimensions (Vitality, Norm, and Modesty) and eight main categories, indicating that theoretical saturation was achieved.

Based on this, a Li-based driving yielding trust mechanism model is proposed (Figure 2), illustrating how situational factors, behavioral expression, and trust formation interact within the framework.

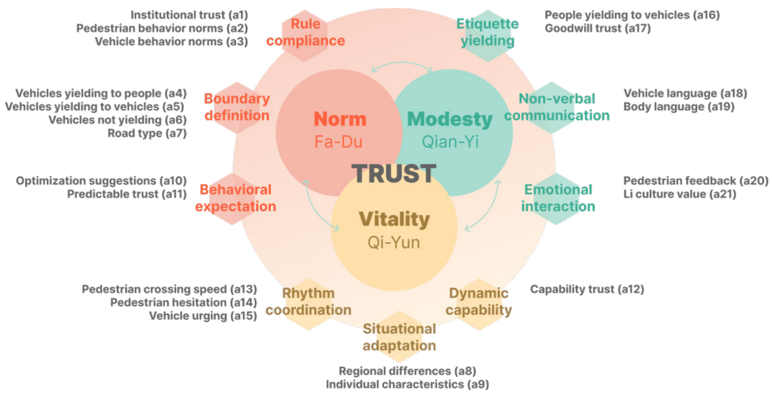

Figure 2: Interaction etiquette-aesthetics model.

AI-EMPOWERED INTERACTION DESIGN FRAMEWORK

"AI Technology-Aesthetic Dimension" Dual-Dimensional Adaptation Framework

Based on the "Interaction Etiquette-Aesthetics" fusion model and considering the unique capabilities of AI technology, this study constructs a dual-dimensional adaptation framework. This framework clarifies the precise correspondence between AI technology layers and interaction aesthetic dimensions: the Perception Layer focuses on the personalized expression of Emotional Aesthetics (Modesty); the Decision Layer emphasizes the scenario-based adaptation of Orderly Aesthetics (Norm); and the Execution Layer centers on the smooth presentation of Rhythmic Aesthetics (Vitality). This correspondence forms the core logic of "Technology supports Aesthetics, Aesthetics optimizes Interaction, and Interaction enhances Trust" (see Table 2).

Table 2: "AI Technology-aesthetic dimension" dual-dimensional adaptation table.

Etiquette Dimension	Aesthetic Representation	Corresponding AI Layer	Technical Implementation & Aesthetic Translation
Vitality	Rhythmic Aesthetics Fluency, Coherence, Aliveness	Execution Layer (Action Control)	Kinematic Smoothing Algorithms: AI adjusts deceleration curves in real-time based on pedestrian walking speed to avoid abrupt braking. eHMI light effects use "breathing" rhythms (e.g., gradient frequencies) to convey the vehicle's "attentive state."
Norm	Orderly Aesthetics Clarity, Certainty, Standard	Decision Layer (Planning & Decision)	Game-Theoretic Decision Models: AI identifies spatial boundaries (crosswalks, traffic lights) and displays determinate "Right-of-Way" signals via projection or screens. Rule libraries ensure consistent yielding logic across different cities (e.g., Hangzhou vs. others).
Modesty	Emotional Aesthetics Non-intrusiveness, Goodwill, Propriety	Perception Layer (Perception & Understanding)	Intent Prediction & Affective Computing: AI identifies pedestrian hesitation, waving, or eye contact. eHMI employs desaturated colors and soft graphics to avoid the oppression of excessive alerts, realizing an interaction atmosphere where "Harmony is Precious."

Layered Implementation Strategy for Interaction Aesthetics

Explicit Norm: "Rule" Expression Based on AI Spatial Cognition

"Norm" is the baseline for establishing trust. AI precisely defines interaction boundaries through high-definition maps and sensor fusion. The design should adopt a Graded Salience Strategy: before a pedestrian enters a conflict zone, the eHMI conveys a clear "Please Cross" signal through high-contrast graphics (e.g., ground-projected crosswalks); during stable driving, signals remain in a low-energy, normalized display mode to ensure a sense of logical certainty.

Implicit Vitality: "Spirit" Simulation Based on Motion Prediction

"Vitality" addresses the mechanical coldness. AI predicts the pedestrian's path for the next 3-5 seconds and aestheticizes the vehicle's braking process. For example, the fluctuation frequency of the front light band can form a physical "resonance" with the pedestrian's step frequency. This dynamic synergy allows pedestrians to perceive the vehicle not as a cold machine, but as an entity with consciousness and care.

Moderate Modesty: "Propriety" Control Based on Context Perception

"Modesty" aims to eliminate AI aggression. Utilizing AI emotion recognition technology, when the vehicle detects signs of anxiety or hesitation in pedestrians (such as the elderly or children), it proactively increases the stopping distance and displays more affable visual symbols (such as soft warm tones) through the eHMI. This expression of "taking a step back" is the digital embodiment of the humble posture in Chinese "Li" culture.

CONCLUSION

As vehicles evolve into intelligent social agents, their design must transcend functional safety to embrace social acceptability.

This study makes three key contributions. First, it operationalizes the concept of Li into three actionable interaction dimensions—Vitality, Norm, and Modesty—providing a measurable cultural lens for vehicle–pedestrian interaction. Second, it develops an Interaction Etiquette Model that explains how culturally embedded expectations shape pedestrian trust calibration in ambiguous traffic scenarios. Third, it proposes an AI–Aesthetics dual-driven framework that bridges algorithmic decision-making and socially legible interaction expression, offering practical guidance for transforming machine intent into emotionally reassuring cues.

Together, these contributions extend existing eHMI research beyond functional signaling toward socially adaptive and culturally embedded interaction design, supporting the development of more trustworthy autonomous mobility systems. This work further highlights the importance of integrating cultural intelligence into autonomous systems, positioning eHMI design as a form of social communication rather than mere information display.

Nevertheless, this study has several limitations. First, the analysis is primarily based on online textual data, which may lack experimental validation. Second, the current framework focuses mainly on visual interaction, without fully addressing multimodal communication or real-world system implementation.

Future research should further validate the proposed model through controlled experiments and real-world studies, and extend the framework to auditory and multimodal interaction design, enabling more comprehensive and context-aware communication between autonomous vehicles and pedestrians.

ACKNOWLEDGMENT

The authors thank the anonymous reviewers for their constructive comments, which significantly improved this manuscript.

REFERENCES

Channamallu, S.S., Pamidimukkala, A., Kermanshachi, S., Rosenberger, J.M., Hladik, G., 2025. Understanding user satisfaction with university parking: A grounded theory approach. J. Urban Mobil. 8, 100136. https://doi.org/10.1016/j.urbmob.2025.100136

Colley, M., Rukzio, E., 2020. A design space for external communication of autonomous vehicles, in: 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. pp. 212–222.

Eisele, D., Kraus, J., Schlemer, M.M., Petzoldt, T., 2025. Should automated vehicles communicate their state or intent? Effects of eHMI activations and non-activations on pedestrians' trust formation and crossing behavior. Multimed. Tools Appl. 84, 24405–24429.

Hensch, A.-C., Kreißig, I., Beggiato, M., Krems, J.F., 2022. The effect of eHMI malfunctions on younger and elderly pedestrians' trust and acceptance of automated vehicle communication signals. Front. Psychol. 13, 866475.

Hoff, K.A., Bashir, M., 2015. Trust in automation: Integrating empirical evidence on factors that influence trust. Hum. Factors 57, 407–434.

Islam, M., She, J., 2025. A Preliminary Exploration to Enhance Appropriate Level of Trust in Autonomous Vehicle and Pedestrian Interaction, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications Sage CA: Los Angeles, CA, pp. 1510–1514.

Kalda, K., Pizzagalli, S.-L., Soe, R.-M., Sell, R., Bellone, M., 2022. Language of Driving for Autonomous Vehicles. Appl. Sci. 12, 5406. https://doi.org/10.3390/app12115406

Lian, Y., Zhang, G., Lee, J., Huang, H., 2020. Review on big data applications in safety research of intelligent transportation systems and connected/automated vehicles. Accid. Anal. Prev. 146, 105711. https://doi.org/10.1016/j.aap.2020.105711

Mahadevan, K., Somanath, S., Sharlin, E., 2018. Communicating awareness and intent in autonomous vehicle-pedestrian interaction, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–12.

Mahmood, A., Szabolcsi, R., 2025. A Systematic Review on Risk Management and Enhancing Reliability in Autonomous Vehicles. Machines 13, 646.

Pipkorn, L., Domeyer, J., Mehler, B., Reimer, B., Gershon, P., 2025. Decoding the silent dialogue: Unveiling driver-pedestrian communication dynamics with a hidden Markov model. Transp. Res. Part F Traffic Psychol. Behav. 109, 965–976. https://doi.org/10.1016/j.trf.2025.01.011

Rasouli, A., Tsotsos, J.K., 2019. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. IEEE Trans. Intell. Transp. Syst. 21, 900–918.

Rouchitsas, A., Alm, H., 2019. External human–machine interfaces for autonomous vehicle-to-pedestrian communication: A review of empirical work. Front. Psychol. 10, 2757.

Sosik-Filipiak, K., Osypchuk, O., 2023. Identification of Solutions for Vulnerable Road Users Safety in Urban Transport Systems: Grounded Theory Research. Sustainability 15, 10568. https://doi.org/10.3390/su151310568

Stampf, A., Colley, M., Rukzio, E., 2022. Towards Implicit Interaction in Highly Automated Vehicles-A Systematic Literature Review. Proc. ACM Hum.-Comput. Interact. 6, 1–21.

Sudhakaran, G.M., Booth, C.A., Prabhakaran, A., Abbey, S.J., Georgakis, P., Renukappa, S., Suresh, S., Magoutas, B., Georgiou, M., Karvouniari-Matzakou, D., 2025. Road safety of micro-mobility and vulnerable road users: Insights and issues from a world café. Int. J. Sustain. Transp. 1–25.
Sujon, M., Dai, F., 2021. Social Media Mining for Understanding Traffic Safety Culture in Washington State Using Twitter Data. J. Comput. Civ. Eng. 35, 04020059. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000943
Tabone, W., De Winter, J., Ackermann, C., Bärgman, J., Baumann, M., Deb, S., Emmenegger, C., Habibovic, A., Hagenzieker, M., Hancock, P.A., 2021. Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. Transp. Res. Interdiscip. Perspect. 9, 100293.
Vafaei-Zadeh, A., Jing Yi, T., Hanifah, H., Nikbin, D., Shojaei, S.A., 2025. Examining autonomous vehicle adoption: A media-based perception and adoption model. Travel Behav. Soc. 40, 101041. https://doi.org/10.1016/j.tbs.2025.101041
Xin, X., Jia, N., Ling, S., 2025. Social norms at crosswalks: Impact on pedestrian-yielding behavior and nudge intervention effectiveness. Transp. Res. Part Policy Pract. 192, 104342. https://doi.org/10.1016/j.tra.2024.104342