Advances in Human Factors of Transportation, Vol. 186, 2025, 379–389
https://doi.org/10.54941/ahfe1006527

Exploring the Role of Predictability in Fostering Passenger Trust in Autonomous Ride-Hailing: A Case Study of Apollo Go

Longyu Yuan

Wuhan University of Technology, Wuhan, Luoshi Road 122, China

ABSTRACT

This study examines the impact of different user interfaces and predictable ways of disclosing decision-making processes on passenger satisfaction and trust in an autonomous driving environment. By designing an experiment that included natural and virtual interfaces, as well as open and closed decision-making processes, we assessed passengers' cognitive load, trust, and satisfaction under different interface conditions. Quantitative results indicated that the Natural Mapping Interface (NMI) significantly reduced cognitive load compared to virtual interfaces (Group B: $M = 5.63$ vs. Group A: $M = 3.95$, $p = 0.026$), supporting Hypothesis H1. Additionally, Group C (NMI with auditory disclosure) showed further reductions in cognitive load ($M = 3.53$ vs. Group B: $M = 5.63$, $p = 0.011$), validating Hypothesis H2. Passenger trust was further enhanced when the decision-making process was disclosed through audio cues. NMI with auditory cues was perceived as the most trust-enhancing design. The results of this study provide insights into HMIs that enhance user trust and satisfaction.

Keywords: Autonomous vehicles, AVs, User trust, Predictive transparency, Ride-hailing services

INTRODUCTION

Autonomous vehicles (AV) have attracted significant attention as a disruptive innovation in the transportation sector. Goodall (2014) defined an AV as one capable of operating without human intervention. This capability enables its integration into shared mobility platforms (e.g. Apollo Go). Despite the many benefits of AV, a major barrier to their widespread adoption remains: building passenger trust.

Passenger trust in AV is a key factor in their widespread adoption. Several factors contribute to passenger distrust, including safety concerns, lack of human supervision, and unpredictability of vehicle behavior. Research has shown that trust in autonomous systems is influenced by their safety, reliability, and predictability (Lee & See, 2004; Goodall, 2014). Of these, predictability plays an important role in fostering trust.

When passengers can predict how an AV will behave, they tend to feel safer and less anxious. Predictability helps reduce uncertainty, thereby increasing passengers' confidence in the system. It also aligns the vehicle's behaviour with passengers' expectations, reducing anxiety (Lee & See, 2004). Additionally, predictability enhances passengers' sense of control and makes the vehicle's actions more congruent with their mental model of how the vehicle should behave (Sheridan, 2016).

Improving predictability in AV is crucial for building user trust. Predictable systems reduce cognitive load on passengers, thereby making them feel more relaxed and confident in the technology.

Transparency also plays a key role in fostering trust in AV. These vehicles rely on artificial intelligence (AI) to sense their environment and make decisions. However, the complexity and invisibility of these processes can lead to a lack of understanding, which may increase passenger anxiety. The "black box" effect occurs when passengers cannot comprehend the reasons behind the vehicle's behaviour.

User interface (UI) design critically shapes passengers' perception of predictability and transparency. A trustworthy UI must present information in an easy-to-understand manner. It should also provide clear feedback on the vehicle's behavior and decision-making processes. The concept of "natural mapping," where the system's behavior aligns with the user's mental model, is key to improving the user experience. By designing interfaces that are intuitive and clear, it is hoped that passenger anxiety can be reduced and trust can be strengthened.

Natural Mapping Theory (NMT) posits that aligning system behavior with user expectations reduces cognitive load by leveraging intuitive design principles (Norman, 2013). Empirical studies have demonstrated that interfaces mimicking real-world interactions enhance usability and lower anxiety. For example, Lee and See (2004) found that predictable automation reduces psychological stress and fosters user trust. Supporting this, Norman (2013) emphasized that natural mappings between controls and outcomes are critical for minimizing learning effort, while Endsley (1995) highlighted the role of situational awareness in ensuring users can anticipate system actions. Such consistency accelerates user adaptation, thereby improving both ease of use and overall experience.

This study aims to improve passengers' satisfaction in autonomous ride-hailing services through passenger-facing UI design. The factor of transparency and predictability of AV and natural mapping principle would be considered to adapt to this paper.

RELATED WORK

Model of Transparency

In response to the transparency challenge, several studies have proposed hierarchical transparency models to better design systems that meet user needs. Lee and See (2004) introduced a basic transparency framework, categorizing it into behavioral and decision-making transparency. Building on this, Chen et al. (2018) proposed the Situation Awareness-Based Agent Transparency (SAT) model to enhance human operators' understanding of autonomous agents' behaviors and states, thereby improving human-autonomy teaming effectiveness.

Automated Driving and Shared Mobility Platforms

With the rise of shared mobility platforms like Apollo Go, autonomous technology is beginning to integrate with shared mobility services. Enhancing passenger trust and comfort in this context has become an important research direction. Existing studies have primarily focused on improving the system interface to build trust. For example, Atakishiyev et al. (2024) suggested that improving the transparency of vehicle behavior in AV can enhance user trust and situational awareness. Providing higher levels of transparency in complex traffic situations helps users better understand the vehicle's decisions and future actions.

Gaps and Challenges in Existing Research

While existing research offers valuable insights into improving the transparency and predictability of AV systems, there are notable gaps. First, most studies have concentrated on behavioral and decision-making transparency, with limited exploration of predictive transparency. Second, many studies are based on simulation experiments or limited user feedback, lacking support from large-scale, real-world data. Therefore, the innovation of this study is to explore how to improve passenger trust in the AV ride-hailing platform in complex traffic scenarios through a hierarchical transparency design based on NMT.

Summary and Resulting Research Questions

Design Problem: In what ways may passenger-oriented central control user interface designs improve user experience and foster trust in AV during rides?

Based on current work, the author identifies two pivotal elements affecting passenger trust in autonomous vehicles: decision transparency (the system's capacity to elucidate its actions) and natural mapping (the correspondence between interface behavior and user cognitive models). Previous research (Lee & See, 2004) highlights that transparency mitigates the "black box" phenomenon, whereas natural mapping alleviates cognitive burden. Nonetheless, scant empirical research has examined the interaction of these aspects inside actual autonomous vehicle interfaces, especially in ride-hailing scenarios such as Apollo Go. To bridge this gap, we devised an experiment to assess the influence of natural mapping and decision transparency on passenger trust. This study seeks to isolate these elements to deliver practical insights for the design of intuitive autonomous vehicle interfaces that enhance user confidence.

RESEARCH HYPOTHESES

To explore the effects of different user interfaces and decision-making process disclosures on passenger satisfaction and trust in an AV environment, we propose the following research hypotheses:

H1: The application of NMT to the user interface can improve passenger satisfaction and performance.

H2: Decision-making process disclosure improves passenger satisfaction and performance.

METHOD

Participants

Since Apollo Go is primarily used by younger demographics, we recruited participants within this group. A total of 19 participants (15 male, 4 female) aged between 18 and 34 years were recruited for the formal experiment, conducted in a semi-open driving simulator. All 19 participants' data were included in the analyses. To minimize potential biases, we ensured that participants had varying levels of familiarity with AV technology, controlling for prior exposure through a pre-experiment questionnaire.

Experimental Design

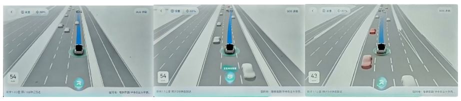
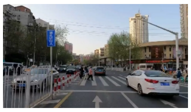
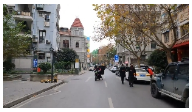
To investigate our hypotheses, the author designed a 2 (interface: virtual/natural) × 2 (decision-making process: public/non-public) within-subjects study. However, given that Apollo Go served as the case study, the author excluded the virtual × decision-making process disclosed condition to reflect real-world constraints, where commercial ride-hailing services typically do not expose decision-making processes through virtual interfaces. Thus, three experimental groups were established (Table 1): Group A: Virtual interface × Decision-making process not disclosed. Group B: Natural Mapping interface × Decision-making process not disclosed. Group C: Natural Mapping interface × Decision-making process disclosed. The decision-making process disclosure in Group C was implemented using auditory cues. Participants were randomly assigned to one of the three conditions to control for order effects. We assessed each test group's cognitive load and interface satisfaction using the NASA Task Load Index (NASA-TLX) scale. the author employed t-tests to analyze the results. Effect sizes (Cohen's d) were calculated to quantify differences. After finishing tasks, we invited participants to our structured interview. Quantitative data were processed by Latent Dirichlet Allocation (LDA) to explore the potential themes.

Materials and Apparatus

These experiments were conducted by using a static driving simulator equipped with a car seat. The design of the driving simulator replicates the operating environment of AV. During testing, participants assumed the role of passengers (see Figure 1). Driving scenes were displayed on a 15-inch screen with a resolution of 1920 × 1080, while an iPad 8 simulated the central control interface. Two independent variables were defined: whether the interface was natural or virtual, and whether the decision-making process was disclosed or undisclosed. All three interface sets adhered to the same information architecture. The Natural Mapping Interface (NMI) was designed to resemble real-world driving conditions, utilizing environmental cues and a first-person perspective to enhance immersion. The virtual interface, in contrast, used a stylized top-down navigation view, focusing on abstract representations of vehicle movement. Auditory cues included pre-recorded phrases such as 'Turning left to avoid pedestrians' and 'Slowing down due to traffic ahead'.

Figure 1: Driving simulators and experimental apparatus.

Table 1: Screenshots of driving scenarios and interfaces under experimental conditions.

Group	Driving Scenario	Central Control Interface
Control Group A		 Virtual + non-disclosure
Experimental Group B		 Natural + non-disclosure
Experimental Group C		 Natural + Disclosure

Interview Procedure and Analysis

In addition to the quantitative experiment, the author conducted a structured interview with all 19 participants after the experiment. One participant's data was excluded due to misunderstanding the AV context. The interview was designed to explore subjective experiences and attitudes toward the three interface conditions (A, B, and C). Each participant was asked the same set of open-ended questions, covering their preferences, interpretation of sound cues, interface comprehension, and perceived trust in the system. The interviews were recorded, transcribed, and subsequently analyzed using

Latent Dirichlet Allocation (LDA) topic modeling to identify major themes. Responses were categorized into three thematic areas: interface, sound, and preferences. To ensure a comprehensive understanding of the experimental findings, the next section will discuss the results and their implications.

RESULT

Descriptive Analysis

The NASA-TLX questionnaire was adapted to focus on three core metrics: Mental Demand, Performance, and Easy-Learning, aligning with the study's emphasis on cognitive load and usability. Descriptive statistics for each experimental group are summarized in Tables 2–4. Group C (NMI + Auditory Disclosure) reported the lowest mental demand ($M = 3.53$, $SD = 2.48$), while Group B (NMI without Disclosure) scored highest ($M = 5.63$, $SD = 2.39$). Group A (Virtual Interface + non-disclosure) showed moderate mental demand ($M = 3.95$, $SD = 2.07$). Group C achieved the highest performance scores ($M = 8.11$, $SD = 2.31$), followed by Group A ($M = 7.89$, $SD = 2.02$) and Group B ($M = 7.47$, $SD = 2.52$). Group C's interface was rated as the easiest to learn ($M = 8.32$, $SD = 1.52$), surpassing Group B ($M = 8.05$, $SD = 1.43$). However, Group A's data ($M = 7.79$, $SD = 2.44$) showed significant outliers, indicating extremely unstable user performance. Consequently, Group A was excluded from further analysis of Easy-learning, and only Groups B and C were compared.

The t-test further validated the differences among the groups. The comparison between Group A and Group B yielded a p-value of 0.026 (< 0.05), indicating that NMI significantly reduce cognitive load, thus supporting Hypothesis H1. A further comparison between Groups B and C showed a significant difference (p-value $= 0.011 < 0.05$), suggesting that the auditory decision disclosure process amplifies the effect, supporting Hypothesis H2. The significant difference between Groups B and C indicates that adding auditory cues enhances predictability, which in turn reduces cognitive load and enhances trust. This result highlights the importance of incorporating auditory cues in user interfaces to improve trust in AV. In summary, the NMI effectively reduces cognitive load, while the combination of auditory decision disclosure (Group C) achieves the optimal effect, verifying the cumulative effect of H1 and H2.

Table 2: Questionnaire design (based on NASA-TLX) control Group A.

Task Load Index (11-Point Likert Scale)		Mean	SD
Mental Demand	How mentally demanding was the task?	3.95	2.07
Performance	Maintenance/installation time to complete	7.89	2.02
Easy-learning	Which interface is easier to learn and understand?	7.79	2.44

Table 3: Questionnaire design (based on NASA-TLX) experimental Group B.

Task Load Index (11-Point Likert Scale)		Mean	SD
Mental Demand	How mentally demanding was the task?	5.63	2.39
Performance	Maintenance/installation time to complete	7.47	2.52
Easy-learning	Which interface is easier to learn and understand?	8.05	1.43

Table 4: Questionnaire design (based on NASA-TLX) experimental Group C.

Task Load Index (11-Point Likert Scale)		Mean	SD
Mental Demand	How mentally demanding was the task?	3.53	2.48
Performance	Maintenance/installation time to complete	8.11	2.31
Easy-learning	Which interface is easier to learn and understand?	8.32	1.52

Correlation Analysis

To illustrate the differences in performance and Easy-learning across the three groups, we generated a box plot of the scores (Figure 2). Cohen's d effect sizes were calculated to assess performance variations between groups, providing insights into passenger trust and satisfaction across different interface designs.

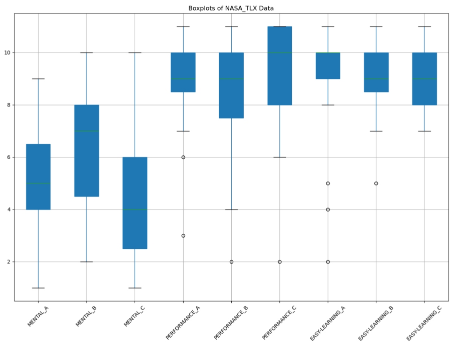


Figure 2: Boxplots of NASA-TLX scores across groups.

In the performance evaluation, Group C exhibited the highest mean scores and was identified as the best-performing group. Group B demonstrated a slight positive effect ($d = 0.18$) compared to Group A, indicating marginally better performance. The comparison between Group A and Group C showed a negligible effect size ($d = -0.10$), suggesting no substantial difference. Meanwhile, the comparison between Group B and Group C revealed A small negative effect ($d = -0.26$) suggested marginally lower performance in Group B compared to Group C. Notably, the absolute effect size was the largest between Groups B and C ($d = -0.26$), highlighting a more pronounced difference. While the direct comparison between Groups B and C did not fully align with expectations, the overall trend suggests that Group B outperformed Group A in performance metrics. Group C's superior performance supports H1 and H2, demonstrating measurable improvements in Groups B and C.

For Easy-learning, Group A exhibited the highest mean; however, the presence of several outliers indicated that user performance was highly inconsistent. As a result, data from Group A were excluded from further analysis. Although Groups B and C had identical mean values, Group C demonstrated a slight positive effect ($d = 0.20$) compared to Group B, suggesting improved learnability. Additionally, Group C exhibited a moderate negative effect ($d = -0.27$) compared to Group A, further underscoring its superiority. The absolute effect size was largest between Group C and Group A ($d = -0.27$), indicating a notable effect. These results strongly validate H2, demonstrating that Group C's transparency-driven design consistently outperformed Groups A and B in terms of Easy-learning. This reinforces the idea that clear decision-making processes enhance user confidence.

Qualitative Interview Results

Participants assigned the highest rating to Group C's interface ($M = 3.91/5$), characterizing it as "intuitive" and "congruent with real-world expectations" (e.g. "*The natural interface is somewhat clearer*"). Group B attained inferior results ($M = 3.22/5$), with critiques emphasizing restricted situational awareness (e.g. "*He does not provide me with a more straightforward, concise perspective*"). Auditory cues in Group C were seen as essential for transparency ($M = 3.91/5$). Participants highlighted that statements such as "*Turning left to avoid pedestrians*" assisted them in "*anticipating the vehicle's subsequent action*" (e.g. "*I must understand my movement*"). Likert ratings corresponded with NASA-TLX outcomes: Group C's superior performance (Table 4) was associated with favorable qualitative feedback, whereas Group B's cognitive load difficulties aligned with interview criticisms.

Table 5: Interview theme analysis and explications.

Theme	Type	Topic	Comment Sample
Interface Design	T1	Insufficient interface realism and operational guidance	The participant found the interface inauthentic. The participant felt the system lacked a simplified view. The participant found the available information limited. The participant noted unnecessary information was filtered out.
	T2	Field of view limitations and information presentation flaws lead to safety hazards	
	T3	Information filtering and support function effectiveness	

Continued

Table 5: Continued

Theme	Type	Topic	Comment Sample
Sound Feedback	T1	Vehicle decision transparency and the need for behavioural interpretation	The participant suggested addressing initial context first. The participant emphasized understanding movement. The participant noted people tend to rely on audio cues. The participant highlighted rear passengers' flexibility in switching.
	T2	Personalization and scene adaptation of sound cues	
	T3	Intelligence and user adaptability of sound alerts	
User Preferences	T1	The need to balance vision and security	The participant suggested visual labelling for clarity. The participant felt auditory cues are necessary. The participant found the virtual interface could be refined. The participant noted the natural interface was clearer.
	T2	Prioritizing comfort and natural interaction	
	T3	Interface adaptation and information clarity in complex scenarios	

DISCUSSION

This study is an early empirical application of NMI in a public automated ridesharing environment, an area that has received limited academic attention compared to private or driver-oriented automated systems. While previous studies, such as Chen et al. (2018), have focused on unimodal transparency mechanisms, such as visual-only or auditory-only disclosures, this study systematically investigated both unimodal and multimodal (coupled) human factors, thus contributing to a more comprehensive understanding of the transparency of the automated driving passenger interface.

The results of the study indicated that the multimodal coupling implemented in Group C (natural interface with auditory decision disclosure) consistently outperformed Groups A and B on a number of NASA-TLX metrics, particularly in terms of performance and ease of learning. This finding supports the hypothesis that integrating visual and auditory channels creates a synergistic effect, leading to a greater degree of user trust and system usability. For example, Group B (natural interface without auditory cues) showed a slight improvement over Group A (virtual interface), but Group C demonstrated a much greater positive response, confirming that the integration of sensory modalities enhances passengers' understanding and confidence.

These results corroborate previous research that has emphasized the role of transparency in improving perceptions of safety and system reliability (Lee & See, 2004). However, unlike these previous studies that focused on drivers or private environments, this study emphasizes that passengers who lack control over vehicle operations may benefit more from multimodal coupled transparency.

In addition to enhancing trust, the use of NMI in public-facing self-driving cars opens up promising directions for user-centric personalized services. Feedback from participants indicated that they desired personalized interaction mechanisms. In the future, AV systems might allow users to choose different levels of transparency, letting them change how much information they receive (e.g. just visuals or both visuals and sounds) to fit their individual thinking styles and needs. This concept of a customizable, gradient-transparent interface offers a novel framework for the inclusive design of automated transportation.

This study has certain limitations. The sample exhibited a gender imbalance and was predominantly composed of a younger demographic, potentially restricting its generalizability. The experimental design utilized a static driving simulator, which, while realistic, fails to fully capture the complexities of real-world driving scenarios. Third, the test scenarios primarily covered standard urban traffic conditions, which are consistent with the environment in which Apollo Go operates in Wuhan, but excluded rare or emergency situations (e.g. bad weather or accidents). Finally, while multimodal interfaces improve accessibility for the average user, they are still limited for people with vision or hearing impairments. Future iterations should explore tactile or olfactory cues to support multisensory inclusion.

Despite these limitations, the preliminary results offer valuable insights for the design of future AV interfaces. The study presents novel concepts and viewpoints for the design of AV interfaces. In the context of the gradual penetration of AV technology into public services, the findings of this study can provide a scientific basis for designing interfaces that are more intuitive, easy to accept, and enhance user trust.

CONCLUSION

In this study, the author explore the effects of different interface types and decision disclosure methods on passenger cognitive load, trust, and satisfaction in an AV environment. Some hypothesis were supported. NMI (especially those incorporating the decision disclosure process) can significantly reduce cognitive load on passengers and increase their trust in the AV system. In addition, we found that passenger satisfaction with the interface was related to the cognitive load and trust that the system elicited. NMI with auditory cues was perceived as the most trust-enhancing design. NMI should prioritize visual-audio synchronization to enhance transparency and improve user trust.

At the same time, the partial support for H1 in quantitative data ($p = 0.026$) aligns with qualitative feedback, where participants perceived NMI as more intuitive despite prior familiarity with virtual interfaces. However, qualitative feedback highlighted that NMI aligned better with their mental models, suggesting a potential gap between subjective and objective measures.

REFERENCES

Atakishiyev, S., Salameh, M., Yao, H. and Goebel, R. (2024) 'Explainable artificial intelligence for autonomous driving: A comprehensive overview and field guide for future research directions', *IEEE Access*.

Chen, J. Y., Lakhmani, S. G., Stowers, K., Selkowitz, A. R., Wright, J. L. and Barnes, M. (2018) 'Situation awareness-based agent transparency and human-autonomy teaming effectiveness', *Theoretical Issues in Ergonomics Science*, 19(3), pp. 259–282.

Endsley, M. R. (1995) 'Toward a theory of situation awareness in dynamic systems', *Human Factors*, 37(1), pp. 32–64.

Goodall, N. J. (2014) 'Machine ethics and automated vehicles', in *Road Vehicle Automation*, pp. 93–102.

Lee, J. D. and See, K. A. (2004) 'Trust in automation: Designing for appropriate reliance', *Human Factors*, 46(1), pp. 50–80.

Norman, D. A. (2013) *The Design of Everyday Things*. New York: Basic Books.

Sheridan, T. B. (2016) 'Human–robot interaction: Status and challenges', *Human Factors*, 58(4), pp. 525–532.