
# Trust Repair of Automated Driving System: A New In-Vehicle Communication Strategy of Voice Assistant

Zhe Song(✉), Ye Yang, and Mengying Deng

Banma Network Technology Co., Ltd, Shanghai 210100, China
songzhe.song@alibaba-inc.com

**Abstract.** Chinese market penetration rate of automated driving systems (ADS) is increasing rapidly. Users are willing to try ADS, but the negative feeling is along with the substantial experience as well. One reason is the gap between users experience and expectation of ADS function, which was formed based on the market information. The other reason is ADS cannot meet different individuals' driving preferences and habits in short term.

As a consequence, users might decrease their trust with ADS, therefore reducing the usage frequency and losing opportunities to rebuild trust. This counteract with the original intention of ADS, which is to improve driving safety.

In the human-machine cooperative ADS, trust repair is necessary for maintaining the trust between the human and the system; in terms of method, anthropomorphic in-vehicle voice communication can enhance the degree of trust. However, there are scarce amounts of studies regarding these two concepts.

Regarding the circumstance, our research proposes a trust repair strategy that is centered on voice communication in ADS. Based on existing ADS technical capabilities, our goal is to improve users' trust and experience with ADS in the early stage of use.

Through market user research, our team systematically summarized the types of scenarios and reasons for the reduction of trust in ADS as the basis of our research. Furthermore, based on the concept of trust repair, a voice-communication-based interaction strategy for ADS is established, and specific dialogues are designed. Finally, a scenario simulated user test (N = 60) was conducted to verify the effectiveness of the strategy: this trust repair approach can significantly improve users' trust in the early use of ADS and their subjective attitudes to use it. Overall, the results provide a new perspective and direct implications for ADS and in-vehicle voice assistant designers.

**Keywords:** Automated Driving · Trust Repair · Voice Assistant

## 1 Introduction

The adoption of Automated Driving System (ADS) is steadily increasing, with sales of Level 2 ADS vehicles in the Chinese market reaching 2.88 million in the first half of 2022, up 46.2% year on year.


C. Stephanidis et al. (Eds.): HCII 2023, CCIS 1836, pp. 99–106, 2023.
https://doi.org/10.1007/978-3-031-36004-6_14

ADS aims to reduce driving stress and enhance safety. However, a recent survey of Chinese consumers found that 39% expressed concerns about how to handle ADS malfunctions, indicating uncertainty in the transition from manual to cooperative driving with [1] ADS. User trust is directly related to automation usage, as shown in a study by Lee and See [2]. Lack of trust in autonomous vehicle systems is a frequently cited reason for driver reluctance to use them, according to recent surveys[3, 4]. Supporting appropriate trust is critical in avoiding misuse and disuse of automation [2, 5, 6]. Therefore, trust is a critical factor in the ADS experience.

Trust consists of initial trust and dynamic trust[7]. A common strategy to increase the likelihood of using ADS is to establish initial trust through pre-teaching or providing guidebooks. Dynamic trust, which changes during ADS driving, is the focus of this study. The study considers how users update their level of trust during the learning phase[8] and how the design of the HMI affects users' dynamic trust and human-machine team performance.

Negative interactions are found to have a greater impact on dynamic trust than positive interactions[9]. During the learning phase, a negative experience can cause a decrease in trust or even lead to users abandoning the use of ADS, as researched by Fredrick Ekman[8]. There are two advantages to dynamically repair trust in the vehicle. Firstly, users can continuously understand the capabilities of ADS by receiving timely information and encountering different situations with relevant explanations. Secondly, ADS can receive real-time data, such as user behavior, emotional state, and environmental changes, and provide more precise explanations and caring communications. To effectively repair users' trust and improve their experience, we propose to implement a voice assistant communication strategy in ADS. Therefore, we first ask the following two general research questions.

**RQ1:** How can trust in ADS be repaired through voice communication?

**RQ2:** How do users' trust and attitudes towards ADS vary with and without trust repair strategies?

## 2 Research

### 2.1 Negative Experience and Reduction of Trust

An in-depth interview was conducted in the early stage of this research with 10 Chinese users. All of them had purchased a new car equipped with an ADS within half a year and had some basic knowledge and experience of the ADS. Based on the interviews, four main factors were identified regarding costly acts and negative experiences:

**The Imperfection of Conditional ADS Functionality.** In some situations, ADS may suddenly exit, and its performance in complex road conditions is unstable.

**Mismatched Driving Style due to Differences in System and Individual Preferences.** In a short period of time, automatic driving functions may not be able to match the driving habits and preferences of different individuals. The limited settings offered by manufacturers may not be adaptable to different dynamic scenarios. The

resulting tense experience may lead users to take over driving, which can reduce their trust in ADS and decrease the frequency of use [10].

**Cognitive Bias and "Disappointment".** Users' initial cognitive understanding of new technologies often comes from the media and may lead to high expectations. During the early stage of usage, understanding and summarizing the system logic and boundaries can be challenging and time-consuming due to the complexity of the ADS function. Additionally, some users may not be interested in the underlying logic. Unchecked high expectations can cause frustration, a decrease in trust [11], and disuse, if the system fails to meet users' expectations during actual experiences [12].

**Lack of Transparency and Communication Channels.** The primary issues are concentrated in the following areas: a lack of understanding of specialized terminology and symbols, unclear identification of the specific reasons for system failures, and an inability to predict system behavior. There are few convenient channels for obtaining accurate information, causing users to rely on community or online searches for help. Furthermore, effectively describing the problems remains a challenging task.

### 2.2 The Main Factors Affecting Trust Repair in ADS

**Beneficial Acts.** According to the Transactional Model of Trust Repair by Ewart, beneficial acts can repair trust, which are perceived as positive or pleasant interactions by the human[13]. For ADS, such beneficial acts may include improved performance, recognition of costly acts, empathetic responses, clear explanations for system's behavior, maintaining a positive tone, demonstrating learning capabilities, and making promises for future improvements[13].

**Timing.** First, humans have learned to trust the system that exhibit expected or predictable behavior[14]. To align users' expectations to a more objective level and reduce feelings of frustration, it can be helpful to inform them of the system's limitations and reasons in advance. Second, when users encounter negative experiences or have inquiries, active communication can be employed to provide timely and accurate responses. Finally, in highly stressful situations, such as sudden takeovers, providing complex explanations may increase safety risks. Therefore, the timing of post-explanation after costly acts is crucial in such situations.

**Voice Communication.** Anthropomorphization is an approach to trust repair [13]. By explaining complex and technical operations in a more accessible way, users are more likely to trust the system. Through interviews, it was found that some users asked an experienced co-pilot about ADS-related questions during the learning phase. The voice assistant can combine vehicle data to support fuzzy search and provide instant feedback, enabling users to ask questions such as "What was that sound just now?" or "What does that icon mean?" Moreover, considering the multitasking in driving, the speech style used in voice assistance needs to be direct and concise.

### 2.3 Communication Strategy of Trust Repair in ADS

In conclusion, a communication strategy for an in-vehicle voice assistant was proposed as a mean of repairing trust in ADS (Fig. 1). This strategy was complemented by a specific conversation design tailored to typical ADS scenarios. The objective of this proposal is to enhance users' trust and overall user experience during the initial phase of use effectively, based on the same technical capabilities. The effectiveness of the repair strategy and dialogue design was validated through comparative testing.

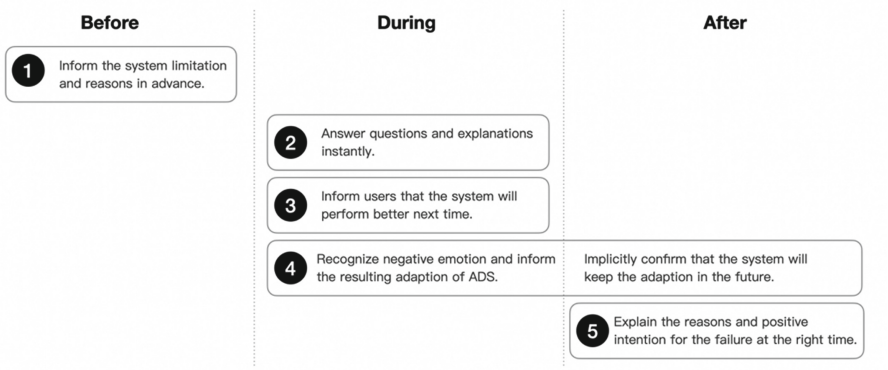


**Fig. 1.** In-vehicle communication strategy for ADS based on trust repair theory.

## 3 Method

### 3.1 Questionnaire

The online survey questionnaire consists of five scenarios, each divided into two groups. One group includes the use of a voice assistant to repair trust (TR), while the other (NTR) does not.

Scenario 1 (S1): Driving home on a rainy night. In the TR, the voice assistant reminds users in advance: "the weather is bad and may affect the performance of the ADS." A message is also displayed on the dashboard when user turns on the ADS. In NTR, there is no voice prompt.
Scenario 2 (S2): An unfamiliar icon appears on the dashboard and the user asks the voice assistant for help. In NTR, the voice assistant will assist in opening the user manual page, while in TR, the assistant will directly explain the meaning of the icon to the user.
Scenario 3 (S3): While driving on an elevated road, if the vehicle system detects that the user is trying to maintain a distance from a large vehicle, the voice assistant in the TR will inform the user that the system has learned about their driving habits and will execute it in the future. On the other hand, the NTR does not have any voice prompts.

Scenario 4 (S4): When the preceding vehicle changes lanes, the automated driving vehicle may quickly accelerate to close the distance, which can cause discomfort for the user due to consecutive acceleration and sudden braking. In both TR and NTR, the system detects the user's discomfort and adjusts the acceleration strategy accordingly. However, in TR, the voice assistant informs the user that it has recognized the discomfort and changed the acceleration mode. After discomfort was reduced, the assistant confirms with user by saying "the same driving mode will be maintained in the future".
Scenario 5 (S5): Due to the inability to accurately recognize unclear lane markings on the elevated road, the ADS exits directly and requires the user to take over the vehicle immediately. Both TR and NTR will display message on the dashboard without any communication (see Fig. 2 a). When the user's cognitive load is reduced or the vehicle comes to a stop, TR's voice assistant provides explanation to user (see Fig. 2 b). NTR does not provide any further communication or explanation.

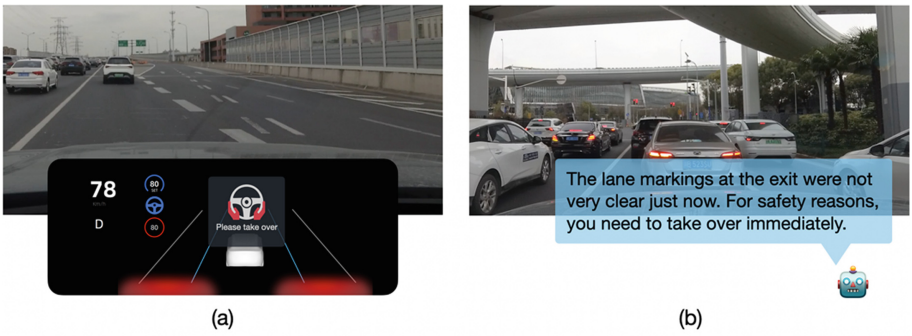


**Fig. 2.** Pictures of scenario 5 in the questionnaire.

### 3.2 Procedure and Measurement

In the questionnaire, the participants were instructed to imagine themselves using a new ADS and experienced both TR and NTR in five different scenarios. The user experience was evaluated by two criteria: "Attitude" and "Trust". The attitude questionnaire was designed based on a 7-point semantic differential scale and divided into 4 dimensions (from 1 to 7 points): "Complicated - Simple", "Obstructive - Supportive", "Foolish - Intelligent" and "Conservative - Creative". The trust scale, adapted from Choi and Ji [15], measured 3 dimensions: TRU1 - "The system is dependable", TRU2 - "The system is reliable" and TRU3 - "Overall, I can trust the system". The questionnaire has distributed a total of 60 copies, with the Cronbach's α coefficient being 0.959, the KMO value being 0.742, and Bartlett's sphere test being $\chi 2=2353.706$, $p < 0.05$, which is indicated to be suitable for analysis.

# 4 Result

## 4.1 The Effect of Trust Repair Strategies on Trust

Table 1 demonstrates that the trust repair strategies (TRS) have a significant impact on the trust of the participants ($p = 0.01 < 0.05$), particularly for the participants over 46 years old. The results indicate that the age of the participant had a significant influence on trust repair ($p = 0.03 < 0.05$).

**Table 1.** Result of Paired Sample T-Test on trust.

| | Groups | Mean | SD | Difference | t | p |
|---|---|---|---|---|---|---|
| Pair1 | TR-TRU1 | 5.45 | 1.59 | 0.47 | 2.087 | 0.041* |
| | NTR-TRU1 | 4.98 | 1.78 | | | |
| Pair 2 | TR-TRU2 | 5.17 | 1.66 | 1.65 | 4.392 | 0.000** |
| | NTR-TRU2 | 3.52 | 1.94 | | | |
| Pair 3 | TR-TRU3 | 5.18 | 1.63 | 0.48 | 2.039 | 0.046* |
| | NTR-TRU3 | 4.7 | 1.87 | | | |

Note. * $p < 0.05$ ** $p < 0.01$

## 4.2 The Effect of Trust Repair Strategies on Attitude

Over all, the participants in the study found that the communication based on TRS was helpful, and the post-explanation voice had a significant impact on their trust in the system. As shown in Table 2, the TRS had a significant impact on the Participant's attitude in S5. Specifically, when the user was required to take over urgently, the post-explanation voice increased their trust in the system.

**Table 2.** Results of Paired Sample T-Test on attitude for 5 scenarios.

| | S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|---|
| Simple | 0.78* (4.93–4.15) | 0.32 (5.2–4.88) | 0.02 (5.25–5.23) | -0.37 (4.8–5.17) | 0.7* (4.75–4.05) |
| Supportive | 0 (4.98–4.98) | 0.17 (5.48–5.32) | 0.37 (5.63–5.27) | 0.33 (5.3–4.97) | 0.75* (5.12–4.37) |
| Intelligent | 0.25 (5.1–4.85) | -0.08 (5.28–5.37) | 0.17 (5.67–5.5) | 0.23* (5.3–5.07) | 0.6* (5.15–4.55) |
| Creative | -0.12 (4.55–4.67) | -0.02 (5.17–5.18) | 0.25 (5.48–5.23) | 0.35 (5.3–4.95) | 0.78** (5.03–4.25) |

Note. The values are: Difference (TRmean - NTRmean) * $p < 0.05$ ** $p < 0.01$

The strategy of enhancing the transparency of ADS to restore trust is effective in improving users' overall evaluation of the system's supportiveness. The results of S1 indicate that the repair strategy significantly improved users' perception of the interaction as being simpler. However, it did not demonstrate a significant difference in creativity and supportiveness compared to NTR. Based on the results of S4, participants felt significantly simpler, but more complicated in TR.

## 5 Conclusion and Discussion

### 5.1 Conclusion

In general, TRS improves users' trust and subjective evaluation in using ADS. Users feel supported in all the scenarios, proving the communication strategy is effective.

The timing of voice assistant communication should take into account the user's real-time emotions and cognitive load, especially in emergency situations such as taking over control (scenario 5). Additionally, it is important to provide explanations for ADS failures, which can help users gain a clear understanding of the system's limits and capabilities.

Directly providing feedback can significantly simplify and enhance the user experience (S1). The voice design strategy should address users' cognitive issues more instinctively to reduce the learning cost. Due to the inability to predict user intents accurately, there was no improvement in the dimensions of intelligence and creativity (S2). S4 made it more complicated for users, possibly because it introduced one more round of dialogue compared to the NTR.

There is also a notable increase in trust and attitude scores towards the ADS, especially among participants aged 46 years old and above.

### 5.2 Limitation and Future Work

The study has several limitations. Firstly, the survey was conducted through online questionnaires, which may not fully represent real-life scenarios. Participants responded to pictures and scenario descriptions, while voice conversations were represented in text form. Future research should incorporate real vehicle environments and live voice assistant conversations. Secondly, measuring initial trust and trust after negative experiences could provide a better evaluation of the effectiveness of the repair strategy. Thirdly, the appropriate level of user trust was not discussed, and communication between ADS and humans should be restrained to avoid over-trust. Lastly, the study found age differences in the effectiveness of trust repair strategies, but future comparative research could investigate the impact of other demographic characteristics such as gender and driving experience.

## References

1. J.D.POWER Webpage article. https://china.jdpower.com/resources/china-self-driving-confidence-index. Accessed 13 May 2021

2. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Human Factors J. Human Factors Ergon. Soc. **46**(1), 50–80 (2004). https://doi.org/10.1518/hfes.46.1.50.30392
3. Hillary, A., Lee, C.: Autonomous Vehicles and Alternatives to Driving: Trust, Preferences, and Effects of Age (2017)
4. Zmud, J., Sener, I.N., Wagner, J.: Self-driving vehicles: determinants of adoption and conditions of usage. Transp. Res. Rec. J. Transp. Res. Board **2565**(1), 57–64 (2016). https://doi.org/10.3141/2565-07
5. Parasurman, R., Riley, V.: Humans and automation: use, misuse, disuse, abuse. Human Factors **39**(2), 230–253 (1997)
6. Berman, S.L., Wicks, A.C., Kotha, S., Jones, T.M.: Does stakeholder orientation matter? The relationship between stakeholder management models and firm financial performance. Acad. Manage. J. **42**(5), 488–506 (1999). https://doi.org/10.2307/256972
7. Hoff, K.A., Bashir, M.: Trust in automation: integrating empirical evidence on factors that influence trust. Hum Factors. **57**, 407–434 (2015). https://doi.org/10.1177/0018720814547570
8. Ekman, F., Johansson, M., Sochor, J.: Creating appropriate trust in auto-mated vehicle systems: a framework for HMI design. IEEE Trans. Human-Mach. Syst. **48**, 95–101 (2018). https://doi.org/10.1109/THMS.2017.2776209
9. Roderick, M.K.: Trust and distrust in organizations: emerging perspectives: enduring questions. Ann. Rev. Psychol. **50**, 569–598 (1999)
10. Huysduynen, H. H., Terken. J., Eggen.B.: Why disable the autopilot? In: 10th International ACM Automotive UI (2018) . https://doi.org/10.1145/3239060.3239063
11. Cohen M.S., Parasuraman R., Freeman.T.J.: Trust in Decision Aids: A Model and Its Training Implications (2002)
12. Dzindolet, M.T., Pierce, L.G., et al.: The perceived utility of human and automated aids in a visual detection task. Hum. Factors **44**(1), 79–94 (2002). https://doi.org/10.1518/0018720024494856
13. de Visser, E.J., Pak, R., Shaw, T.H.: From 'automation' to 'autonomy': the importance of trust repair in human–machine interaction. Ergonomics **61**, 1409–1427 (2018). https://doi.org/10.1080/00140139.2018.1457725
14. Haspiel, J., Du, N., Meyerson, J., et al.: Explanations and expectations: trust building in automated vehicles. In: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2018). Association for Computing Machinery, New York, NY, USA, pp. 119–120 (2018). https://doi.org/10.1145/3173386.3177057
15. Choi, J.K., Ji, Y.G.: Investigating the importance of trust on adopting an autonomous vehicle. Int. J. Human-Comput. Interact. **31**(10), 692–702 (2015). https://doi.org/10.1080/10447318.2015.1070549