An Exploration of Potential Factors Influencing Trust in Automated Vehicles

Hatice Şahin, Sarah Vöge, Benjamin Stahr, Nina Trilck, Susanne Boll

An Exploration of Potential Factors Influencing Trust in Automated Vehicles

Hatice Şahin, Sarah Vöge, Benjamin Stahr, Nina Trilck, and Susanne Boll

University of Oldenburg, Germany firstname.lastname@uni-oldenburg.de

Abstract. Trust, which is one of the main components for acceptance of automated vehicles could be affected by different factors. We have investigated the influence of prior information regarding the safety of Automated Vehicles, different light conditions and the malfunction of external Human-Machine Interfaces on trust. Despite a small sample, we have found that trust is reduced when malfunctions occur and it is immediately recovered back. Prior information regarding safety and different light conditions didn't have a significant effect on trust.

Keywords: Automated Vehicles · eHMI · Overtrust · VR · VRU · Acceptance · Malfunction

1 Introduction and Background

During the last years, an increasing interest in the research of autonomous and automated vehicles (AV) could be observed [9]. Besides the technical challenges, human acceptance is an important aspect deciding about the deployment of AVs. In the public opinion, often only the acceptance by passengers of the AV seems to be important [6]. However, the acceptance by other road users is also important. Therefore, a focus should also be on the perception of vulnerable road users (VRUs) [6]. Especially for VRUs, it is hard to understand the intentions of AVs which leads to acceptance problems [6]. This problem is mainly based on the lack of interpersonal communication via gaze, gestures and/or facial expressions between VRUs and the driver [6]. Hence, the effective communication between VRUs and AVs is still an open challenge [6]. Trust, which is one of the main components for acceptance [11], can be influenced by different aspects [7]. Trust and the level of knowledge on AVs has been found to be linked with acceptance of AVs [11]. In this work, trust is to be understood as the trust towards the correct, error-free behavior of AVs, thereby in the whole system. Not only the absence of trust is problematic, but sometimes also its presence [6]. One reason of the latter is overtrust. Overtrust can be defined as the false estimation of risk when interacting with a machine [10].

For the communication between VRUs and AVs, the different modalities of communication such as body language, auditory, haptic, and the visual modality could be considered [4]. The visual modality has been explored the most often and it has been considered as the most intuitive modality for the majority of potential users. It can include texts, symbols, abstract visual shapes and forms or anthropomorphic elements [4]. One of the ways to use this modality is via external Human-Machine Interfaces (eHMI). eHMIs can be integrated on the surface of the AVs in order to present different types of information [4]. They may contain the intention of the AV, automation state, advice, time-to-cross, situational awareness, danger/safety zone and warnings [4]. Some studies have already explored the idea of using eHMIs for the necessary communication in an ambiguous crossing situation. They have mainly focused on single pedestrians [5, 4, 8, 6]. Some other studies also included malfunctions and/or system errors. For example, Holländer et al. [6] conducted a study concerning overtrust towards AVs in virtual reality. The study showed that the initial trust and perceived safety were negatively affected by a single malfunction, but it recovered fast, which indicates overtrust.

2 User Study

To further investigate the role of trust and overtrust in the interaction between VRUs and AVs, a user study was conducted in a VR environment. Seven participants were recruited (one female), aged between 21-85 ($M = 41$, $SD = 25.5$). Three factors with two levels were tested for their potential influence on VRU's trust on AVs. These factors were

- Prior information regarding the safety of AVs as between subjects variable (positive and negative information) [11].
- Different light conditions as within subject variable (day and night). To our knowledge, this factor hasn't been explored in previous research before.
- Influence of experiencing a collision based on an intentional malfunction of the eHMI as within subject variable (match and mismatch)[6]. Whereas, the collision itself would be a between subject variable.

The task was to cross the road in front of an AV. Participants used bluetooth controllers in order to move the virtual world. The experiment consisted of two blocks with five trials each and a training. Before the start of VR trials, participants received an information sheet regarding the safety of AVs. One group received only positive information while the other one received only negative information. The light condition was changed after the first block and the starting condition was counter-balanced among participants. The malfunction which could lead to a potential collision occurred in the third and the eighth trial.
Participants started with filling out the informed consent and a Demographic Questionnaire [3]. It is part of the PRQF which further contains the Pedestrian Behavior Questionnaire, Pedestrian Receptivity Questionnaire, and a Scenario-based Questionnaire. In order to capture trust and possible changes, Pedestrian Receptivity Questionnaire was used before and after reading the information sheet, and after the VR part. Moreover, for differences in trust, the Scenario-based Questionnaire [3] and the STS-AD [2] were examined. The Scenario-based Questionnaire of two questions were handed out after the information and after

each VR trial block. The STS-AD was implemented in the VR to measure trust after each single trial. It is a set of five items of which one is directly about trust. Additionally, the IPQ [1] was used to measure the sense of presence.

3 Initial Results

Starting with the first factor, the trust items of the Pedestrian Receptivity Questionnaire [3] were used in order to answer whether positive or negative information regarding the safety of AVs affected the trust. For this, the results from the PRQ after reading the information were used. No significant differences in trust between positive and negative information groups were found with Mann-Whitney U test ($p = 0.5$). Scores from STS-AD yielded similar results ($p = 0.45$).

The second factor was explored for the purpose of investigating the effect of different light conditions on trust in AVs. By using trust values from STS-AD, Wilcoxon signed-rank test was performed. The p-value of 0.40 indicated no significant differences between the two light conditions. Further, the Scenario-based questionnaire from the PRQF [3] was analyzed. The first question asks the type of behavior on a crosswalk with an AV approaching. The most participants chose to "wait until AV stops", followed by "wait until AV brakes" and only one participant chose to "hurry to cross". Conducting a Wilcoxon signed-rank test, a p-value of 1 underlines the similar choices for the dark and the daylight conditions. The second question asked about the acceptance of presence of AVs in the participants' area. Wilcoxon signed-rank test indicated no significant differences between different light conditions ($p = 1$).

The third potential factor which could affect trust was the influence of experiencing a collision based on a malfunction. Only a single collision occurred among all trials in all participants. This collision led to a high decrease in trust scores of STS-AD and it was observed that the trust of the affected participant was lower than the median trust of other participants.

To test the effect of malfunctions, a Wilcoxon signed-rank test was performed in STS-AD scores, which indicated a significant difference in trust between trials with and without eHMI malfunctions ($p < 0.001$).

4 Discussion and Future Work

The results of the effect of malfunction in STS-AD trust scores are in line with the results of Holländer et al. [6], where the occurrence of a malfunction directly influenced the trust. Our participants had high trust into the system from the beginning and their trust recovered directly in the next trial after the occurrence of a malfunction as in Holländer et. al. However, in terms of the effect of the collision, we cannot derive a meaningful outcome since we had a small sample. Moreover, two participants had some prior knowledge about automated driving, its current status and possible future developments, which could have influenced the results. The use of a VR environment could have increased situational trust, as users would not be harmed even in the event of a collision. However, this is necessary, because collisions could have and did occur in the experiment due to car malfunction. Furthermore, the VR helped us to provide a controlled setting, so the influence of other potentially trust-influencing factors, such as weather conditions or other pedestrians, could be avoided. Some improvements could be made in the future in the VR itself, such as better integration with the world of questionnaires presented in VR, a more complex crossing task with mixed traffic situations, or adding 6 DoF support to allow participants to walk in reality.

References

1. igroup presence questionnaire (ipq) overview, http://www.igroup.org/pq/ipq/index.php
2. Situational trust scale for automated driving (sts-ad): Development and initial validation. In: 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. pp. 40–47 (2020)
3. Deb, S., Strawderman, L., Carruth, D.W., DuBien, J., Smith, B., Garrison, T.M.: Development and validation of a questionnaire to assess pedestrian receptivity toward fully autonomous vehicles. Transportation research part C: emerging technologies **84**, 178–195 (2017)
4. Dey, D., Habibovic, A., Löcken, A., Wintersberger, P., Pfleging, B., Riener, A., Martens, M., Terken, J.: Taming the ehmi jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. Transportation Research Interdisciplinary Perspectives **7**, 100174 (2020)
5. Faas, S.M., Kao, A.C., Baumann, M.: A longitudinal video study on communicating status and intent for self-driving vehicle–pedestrian interaction. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–14 (2020)
6. Holländer, K., Wintersberger, P., Butz, A.: Overtrust in external cues of automated vehicles: an experimental investigation. In: Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications. pp. 211–221 (2019)
7. Lee, J.D., See, K.A.: Trust in automation: Designing for appropriate reliance. Human factors **46**(1), 50–80 (2004)
8. Otherson, I., Conti-Kufner, A.S., Dietrich, A., Maruhn, P., Bengler, K.: Designing for automated vehicle and pedestrian communication: Perspectives on ehmis from older and younger persons. Proceedings of the Human Factors and Ergonomics Society Europe pp. 135–148 (2018)
9. Tango, F., Montanari, R.: Shaping the drivers' interaction: how the new vehicle systems match the technological requirements and the human needs. Cognition, Technology & Work **8**(3), 215–226 (2006)
10. Wagner, A.R., Borenstein, J., Howard, A.: Overtrust in the robotic age. Communications of the ACM **61**, 22–24 (2018)
11. Ward, C., Raue, M., Lee, C., D'Ambrosio, L., Coughlin, J.F.: Acceptance of automated driving across generations: The role of risk and benefit perception, knowledge, and trust. In: International Conference on Human-Computer Interaction. pp. 254–266. Springer (2017)