The Devil is in the Details - Trust Development During Initial Usage of an Automated Vehicle

Mikael Johansson, Chalmers University of Technology, Sweden, *Johamik@chalmers.se*, **Fredrick Ekman,** Chalmers University of Technology, Sweden

ABSTRACT

Trust is important for adoption of automated vehicles (AVs) and has therefore become a well-researched topic due to the introduction of more and more automation in AVs. Previous research has identified aspects such as user variabilities, the automation itself and context to affect the users' level of trust in AVs. Studies have also shown that trust forms through continuous experience with AVs but can decrease in case of incidents with or perceived failures of the automation. However, only a small portion of the research in this area has focused on individual trust formation and to what extent trust formation processes differs between different individuals. Therefore, this paper investigates how user trust develops during first-time usage of an automated vehicle. Nineteen participants experienced a fully automated vehicle on a test track and information on the participants' trust was collected by interviews and an especially for the experiment developed tool, the 'Trust Curve'. The Trust Curve was used to understand how the user's trust changes over-time during initial usage. The findings identified two main trust formation processes; one Inconsistent process, i.e. trust increases and decreases several times during the interaction with the AV, and one Consistent process, i.e. trust continuously increases/decreases or are unchanged during the interaction with the AV. Thus, findings indicate that users develop trust in different ways which may lead to implications regarding how users obtain a proper level of trust as well as how to design AVs for different groups of users.

Keywords: automated vehicles, trust, individual trust formation, user study.

1. INTRODUCTION

Trust is important for adoption of automated vehicles (AVs) (Buckley, Kaye, & Pradhan, 2018; Ghazizadeh, Lee, & Boyle, 2012), especially in the beginning of the interaction between user and AV. It is a complex phenomenon that is affected by system characteristics, user variabilities, and the context in which the automation is operating (Hoff & Bashir, 2015). System characteristics can be communicated in different ways for example has previous research identified that information in displays (Helldin, Falkman, Riveiro, & Davidsson, 2013) and from the AVs driving behaviour (Ekman, Johansson, Bligård, Karlsson, & Strömberg, 2019) affects the users' level of trust. It has also been recognised that factors such as culture, gender and age affect the users' level of trust in AVs (Hoff & Bashir, 2015).

Furthermore, trust is dynamic, it forms through continuous experience with AVs and stabilizes after time as users build a familiarity with the system (Oliveira, Proctor, Burns, & Birrell, 2019; Yang, Unhelkar, Li, & Shah, 2017). In a study by Beggiato, Pereira, Petzoldt, and Krems (2015) the participants level of trust stabilized after a fifth driving session, which corresponded to 185 km or 3.5 h of driving. This trust development can be aided by information from the AV, both through feedforward provided before the AV

takes actions (Haspiel et al., 2018) or specific feedback that corresponds to the situation (Edelmann, Stümper, & Petzoldt, 2019). However, trust can also decrease in case of incidents with or perceived failures of the automation (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003).

Even though, a large amount of research has focused on users' trust in AVs, only a small portion has focused on individual trust formation and to what extent trust formation processes differ between different individuals. Therefore, this paper investigates how users' trust develops during first-time usage of an automated vehicle, in order to understand different individual trust formation processes.

2 METHOD

2.1 Set-up & Procedure

A Wizard of Oz experiment was conducted on a test course using a remodelled Volvo XC90 operated by a professional test driver (the wizard driver) sitting in the backseat behind concealed driving equipment (i.e. steering wheel, pedals and gear lever). The test course included seven traffic situations, designed to mimic everyday traffic situations, such as overtaking a car or bicyclist and stopping at pedestrian crossings. Nineteen participants experienced two different driving styles, 'Aggressive' and 'Defensive', in the AV they believed to be a fully automated [SAE level 5]. The two AV driving styles were both designed to be experienced as competent, only varying in acceleration, deacceleration, starting and stopping behaviour, distance to objects, gear changing and lane positioning. The participants experienced one of the two driving styles in each of two test runs. Each test run took approximately 15 minutes. For a detailed description over driving styles and set-up see Ekman et al. (2019).

2.2 Test Course

The route driven on the test course consisted of two sections: a city area (with several buildings, covering an area of around 8,000 square metres) and a rural road (approximately 6 km long, normal road standards for bi-directional traffic, allowing speeds of 80 km/h).

2.3 Participants

The 19 participants (10 male and 9 female) was recruited from four different groups: males over/under 30 years and females over/under 30 years, to achieve a balanced gender and age distribution (mean age = 36,7; SD 11.1).

2.4 Data Collection

Different methods were used to collect data during the peri-trial phase (in-car data collection) and the post-trial phase (post-test data collection) to collect the participants' momentaneous trust responses, as well as to allow the participants to reflect further and more deeply on how they experienced the two driving styles when encountering the seven traffic situations. This paper focuses on the data collected from a 'trust curve' and post-trial interviews (post-trial phase). The 'trust curve' (adapted from the UX curve (Kujala, Roto, Vaananen-Vainio-Mattila, Karapanos, & Sinnela, 2011)) was introduced to the participants after each test run to, in hindsight, assess their trust formation. The participants were told to illustrate their trust by drawing a curve representing their level of trust over time during the experience

with the AV and to mark the most prominent situation(s) regarding trust. After, they also got to write down a description of why the trust curve looked as it did. In the post-trial phase, the trust curve was brought up (once again) to stimulate participants to further reflect on and discuss their level of trust in the AV. The in-depth post-trial interviews were recorded and later transcribed in full.

2.5 Analysis

A three-step analysis was conducted. **Step 1**: First, a grouping of trust curves by searching for visual similarities and/or differences between participants' trust curves. When similar groups and sub-groups had been identified, each group were defined and described based on the visual similarities among the trust curves in that group. **Step 2**: Based on the participants' written comments after drawing their respective trust curve, a thematic analysis was conducted to identify descriptions that shared similarities regarding trust. For example, comments that had detailed descriptions of how different situations affected their trust were clustered as 'Situation based judgements' and comments that stated the trust was based more on a holistic experience were clustered as 'Overall judgement'. The clusters were then analysed in reference to the groups of curves in order to see if the patterns were consistent and similar for and between both data sets (curves and comments) and to understand why the curves look as they do. **Step 3**: A second thematic analysis was conducted on the post-trial interview data and compared to the results from Steps 1 and 2, in order to see any similarities and/or discrepancies and to further understand the participants' trust formation.

3 FINDINGS

The trust curves are divided into five sub-groups based on their visual similarities; Fluctuating, Single, Positive, Negative, and Neutral. The sub-groups are further divided into two larger groups, 'Inconsistent' that includes the sub-groups with curves that have a few or several major fluctuations, and 'Consistent' that includes the sub-groups without any major fluctuations. Figure 1 shows the curves clustered together in each group as well as a description of each sub-group's curve characteristics.

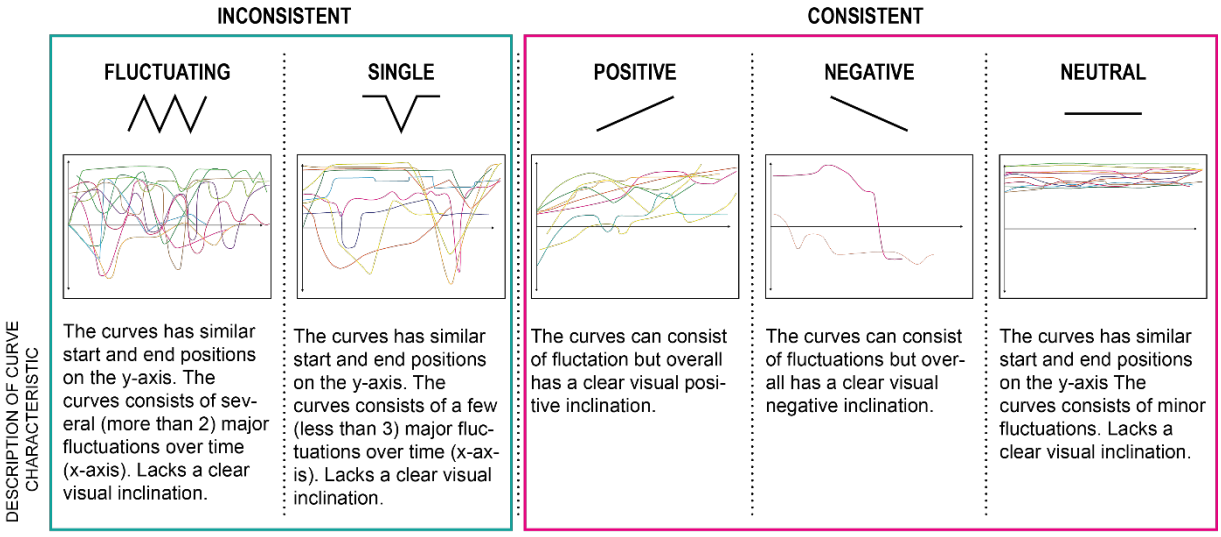

Figure 1. Groups and sub-groups of trust curves, with descriptions of respective sub-groups

curve characteristics

Six of the participants belong to the Inconsistent group (both trust curves drawn by the participant being classified as Inconsistent) while 8 participants belong to the Consistent group (both trust curves drawn by the participant being classified as Consistent) (Figure 2). Thus, participants described their trust formation process in different ways.

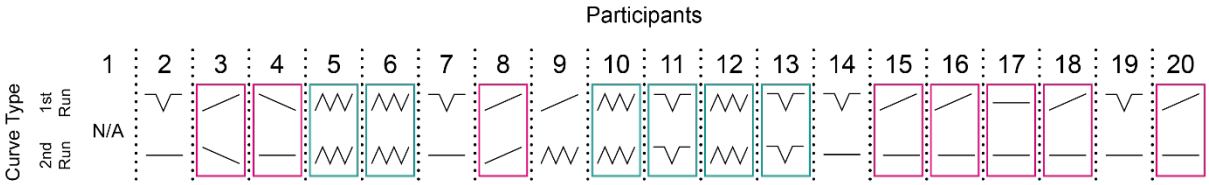

Figure 2. Categorization of the participants' trust curves

The comments made in connection to the trust curves show that the participants belonging to the Consistent group to a higher degree made comments about a holistic experience, focusing on the effect of the overall experience on trust formation rather than the effect of specific situations, (12 vs 6) and more often indicated an incremental rise in the level of trust after each situation (4 vs 1) compared to the Inconsistent group (see Table 1). The interview data followed a similar pattern as the trust curves, and it was evident that some participants had a more holistic way of describing their trust formation processes. One participant stated *"the first test run you are just getting a feel for it [the AV] and there were no problem at all, it was positioned exactly correct and it [trust] only increased all the time"* (P16) not specifying what in the AV driving behaviour or in the traffic situations that affected their trust in the AV. They assessed the interaction with the AV and driving behaviour as an overall experience, and as long the AV managed one situation there was little doubt regarding the following traffic situation.

Table 1. Number of written comments in each category and group.

	OVERALL JUDGEMENT	SITUATION BASED JUDGEMENT	DRIVING BEHAVIOUR BASED JUDGEMENT	INCREMENTAL TRUST FORMATION
INCONSISTENT	6	5	2	1
CONSISTENT	12	4		4

In contrast, participants in the other group, Inconsistent, drew more fluctuating curves, indicating more situational details on the trust curves, and commented more on details of the driving behaviour and situational aspects of the experience. They had a more analytical way of describing their trust formation, that often included a holistic interpretation of the AV but also a situation to situation (and action to action) based way of interpreting if the AV was trustworthy or not. One of these participants explained that "*already when we were about to drive out it* [the AV] *drove all the way to the line* [a line before an

automatic barrier] *and slowed down with a hard brake in the end* [before coming to a full stop]*, that was uncomfortable, but also, didn't it see the line[?], it did not really have any foresight at that moment" (*P10). They analysed the driving style through interpreting the actions the AV did and how the actions were appropriate for different situations or not, more rationally explaining what the AV did, how they felt as well as the possible cause of the specific behaviour.

4 DISCUSSION & CONCLUSION

The findings show that participants describe their trust formation process in different ways. Two different tendencies were evident, one where participants based their trust formation on an overall experience and one where participants based their trust formation on specific behaviors of the AV in relation to the specific situation encountered. It is suggested that the participants whose trust was formed by situation based judgements also analyzed the experience using a more analytical process when assimilating the information (Lee & See, 2004).

Users who, to a higher degree, form trust based on an analysis of AV actions, situation to situation, and continuously evaluating the performance, may therefore not be as susceptible to wrongly calibrated trust as users who judge the interaction with an AV as an overall experience. This since one situation may not be the other alike, so just forming trust on the fact that the AV has performed well in a previous situation or in general and not considering the character of the new situation may cause too high levels of trust in situations that the AV is incapable of handling. Further, Beggiato et al. (2015) showed that information about automation limitations that are not continuously updated tend to disappear from the mental model, which can be more probable for the participants that forms their trust more on the general experience, since they may not be as observant to subtle signs of system limitations. Beggiato et al. (2015) continues by recommending that users should periodically be reminded of system limitations and therefore one could argue that this may be especially important for users that form their trust on the overall experience in comparison to those who already judge the AV trustworthiness based on a situation to situation (and/or action to action) basis.

Hence, the consequence of the difference in trust formation is that users who primarily views the interaction and usage of an AV as an overall experience (without considering situation to situation basis) might form miscalibrated trust, since the devil is in the details.

5 ACKNOWLEDGMENTS

This work was supported by Vinnova, Sweden's Innovation Agency, under grant number 2014-01411. The study was also able to use the facilities and expertise of the AstaZero full-scale test environment through the open research grant, application number A-0025.

REFERENCES

Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation research*

part F: traffic psychology and behaviour, 35, 75-84. doi:https://doi.org/10.1016/j.trf.2015.10.005

Buckley, L., Kaye, S.-A., & Pradhan, A. K. (2018). Psychosocial factors associated with intended use of automated vehicles: A simulated driving study. *Accident Analysis & Prevention, 115*, 202-208. doi:https://doi.org/10.1016/j.aap.2018.03.021

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies, 58*(6), 697-718. doi:10.1016/s1071-5819(03)00038-7

Edelmann, A., Stümper, S., & Petzoldt, T. (2019, 9-12 June 2019). *Specific Feedback Matters - The Role of Specific Feedback in the Development of Trust in Automated Driving Systems.* Paper presented at the 2019 IEEE Intelligent Vehicles Symposium (IV).

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

Ghazizadeh, M., Lee, J. D., & Boyle, L. N. (2012). Extending the Technology Acceptance Model to assess automation. *Cognition Technology & Work, 14*(1), 39-49. doi:10.1007/s10111-011-0194-3

Haspiel, J., Du, N., Meyerson, J., Robert Jr, L. P., Tilbury, D., Yang, X. J., & Pradhan, A. K. (2018). *Explanations and expectations: Trust building in automated vehicles.* Paper presented at the Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction.

Helldin, T., Falkman, G., Riveiro, M., & Davidsson, S. (2013). Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. 210-217. doi:10.1145/2516540.2516554

Hoff, K. A., & Bashir, M. (2015). Trust in automation: integrating empirical evidence on factors that influence trust. *Hum Factors, 57*(3), 407-434. doi:10.1177/0018720814547570

Kujala, S., Roto, V., Vaananen-Vainio-Mattila, K., Karapanos, E., & Sinnela, A. (2011). UX Curve: A method for evaluating long-term user experience. *Interacting with Computers, 23*(5), 473-483. doi:10.1016/j.intcom.2011.06.005

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human factors, 46*(1), 50-80. doi:DOI 10.1518/hfes.46.1.50.30392

Oliveira, L., Proctor, K., Burns, C. G., & Birrell, S. (2019). Driving style: How should an automated vehicle behave? *Information, 10*(6), 219.

Yang, X. J., Unhelkar, V. V., Li, K., & Shah, J. A. (2017, 6-9 March 2017). *Evaluating Effects of User Experience and System Transparency on Trust in Automation.* Paper presented at the 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI).