Understanding Trust in an AV-context: A Mixed Method Approach

Fredrick Ekman, Chalmers University of Technology, Sweden, *ekmanfr@chalmers.se*, **Mikael Johansson,** Chalmers University of Technology, Sweden, I.C. **MariAnne Karlsson,** Chalmers University of Technology, Sweden

ABSTRACT

Trust is a fundamental part of technology acceptance as well as an important factor for creating a positive user experience with Automated Vehicles (AVs). In order to fully understand users' trust in AVs it is important to consider the cognitive processes by which humans develop trust. We argue that a deeper understanding of these processes can be elicited by using a convergent mixed method design. The method design described in this paper was created during an experimental study investigating the effect of AV's driving behaviour on users' trust. The design consists of five data collection methods, three qualitative and two quantitative, used to collect data during and after test runs with an AV. The results show that the different methods elicited responses that may indicate different cognitive processes. The methods used during the test runs produced more affective and analogical responses while the methods used directly after each of the test runs generated more analytic responses. The last method, introduced after the completion of all test runs, produced a more mixed result. The participants elaborated on their earlier responses and sometimes turned their affective responses into analytic or analogical explanations. Hence, by combining and utilizing the strength of different data collection methods, more rich data was elicited on the trust formation process and thereby creating a more nuanced picture of users' trust in automated vehicles.

Keywords: Trust, Human-Machine Interaction, Automated Driving, Automated Vehicles, Mixed-Method Design.

1 BACKGROUND

As technology progresses within the area of semi- and fully automated vehicles, it has become increasingly important to consider and understand trust. It is one of the fundamental factors to create acceptance (Ghazizadeh et al., 2012) as well as a cornerstone to create a positive user experience (Waytz et al., 2014). However, trust is a complex concept making it difficult to assess and available methods are few, especially in the area of automated vehicles (AVs). In order to fully understand users' trust in AVs it is important to understand the cognitive processes that govern trust. It has been suggested that trust involves three different cognitive processes (Lee & See, 2004): (i) the affective process which is the emotional process, i.e. feeling trust; (ii) the analogical process which is connecting earlier and familiar experiences and using them to assess the trustworthiness of an agent (in this case the AV) based on differences and similarities, and finally, (iii) the analytic process which is to rationalize around the agent's trustworthiness. The affective process is the most influential and fundamental trust process for user behaviour, affecting both the analogical and analytic processes of trust. It is also the least cognitively demanding process. The analogical process is used when information about an agent is lacking and earlier experiences (with similar agents) are used to assess the trustworthiness whereas, the analytic process is the logical argumentation of an agent's trustworthiness in which the information about an agent is evaluated (Lee & See, 2004).

With this paper, we argue for a better understanding of drivers'/users' trust in AVs through using a convergent mixed method design (cf. Creswell & Plano Clark, 2017), capturing a more nuanced picture with the aim of exploring the respective cognitive processes of trust in an AV-context.

2 METHOD

The design of the method evaluated was created in connection to an experiment investigating how the behaviour of an AV (in terms of acceleration, deceleration, lane positioning etc.) affects the user's trust. The experiment was conducted on a test course where 18 participants experienced two consecutive AV-test runs with two different driving behaviours. The vehicle had a Wizard of Oz-setup, i.e. it was controlled by a (to the participant) hidden test driver. The test course included a rural- and as well an urban section. Each test run took approx. 15 minutes to complete and included seven 'critical' traffic situations. The seven situations were stopping for a red light, overtaking a moving vehicle, stopping for a person waiting to cross a zebra-crossing, passing a cyclist, driving onto a highway, passing oncoming traffic and driving through a roundabout.

2.1 PARTICIPANTS

The 18 participants taking part in the experiment were 10 men and 8 women between 20 and 55 years old (mean 36,7; SD=11.1). The participants were recruited from the Gothenburg city area (on the west coast of Sweden) through a newspaper advertisement. The only inclusion criterion was a valid driver's license.

2.2 DRIVING BEHAVIOUR

The two distinctly different driving behaviours consisted of a 'defensive' behaviour and an 'aggressive' behaviour. The respective AV-driving behaviours differed in several ways, for example in acceleration, deceleration and distance kept to other objects (Table 1). All participants experienced both behaviours but in different orders, i.e. half of the participants started with the 'defensive' and half started with the 'aggressive' driving behaviour.

Table 1 – AV driving behaviours.

	Defensive Driving Behaviour	**Aggressive Driving Behaviour**
Starting & stopping	Keep the vehicle rolling (avoid standstill)	Start & stop (comes to full stop)
Acc./Decell. Pattern	Avoid heavy acc./decell.	Heavy acc./decell.
Lane positioning	Early indicate right or left turn (through positioning in lane)	Indicate late right or left turn (through positioning in lane)
Distance to objects	Keep longer distance (lateral & longitudinal) to other objects	Keep shorter distance (lateral & longitudinal) to other objects

2.3 PROCEDURE AND DATA COLLECTION

Data collection was performed during two different phases; a peri-trial phase and a post-trial phase (see Figure 1).

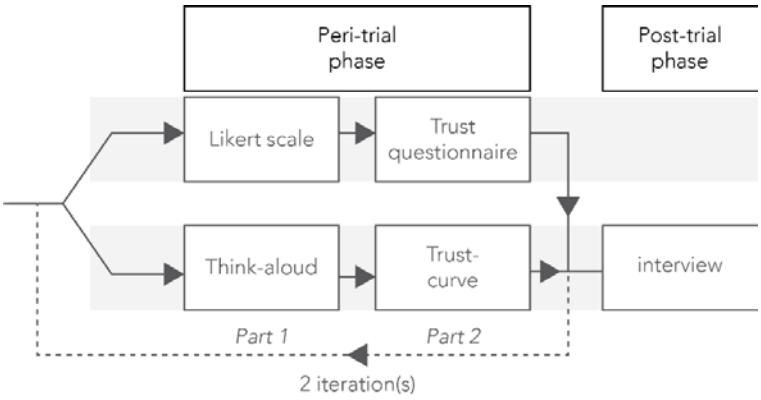

Figure 1 – Data collection methods during peri-trial phase and post-trial phase.

The peri-trial phase included two parts and four data collection methods. The first part was conducted during the test run in the AV, and included a *Likert scale* on trust combined with a *Think-aloud procedure* (cf. Charters, 2003). The participants rated what level of trust they felt in each of the seven critical situations (i.e. passing a cyclist, stopping for a red light, stopping for a pedestrian crossing the street etc.) and then elaborated on why they felt this way. The second part was completed directly after each test run using a *Trust Questionnaire* (based on Jian et al., 2000) and a method referred to as the *'Trust Curve',* an adaptation of the UX-curve (Kujala et al., 2011). The Trust Questionnaire focused on assessing the participant´s overall trust in the AV during each of the respective test runs. The participants were then asked to draw a curve symbolising their level of trust during the test run and to mark out situations that they experienced affected their trust in the AV the most. The post-trial phase, finally, included an interview with the participant, where the 'Trust Curve' was used as a mediating tool to allow the participant to further reflect on and discuss the levels of trust in the AV during specific situations as well as their overall trust.

3 RESULT

Initial analyses show that during the first part of the peri-trial phase, the Likert Scale and the Think-aloud procedure provided more affective and analogical responses from the participants regarding the trustworthiness of the AVs. Some of the affective responses during the test described a general feeling of trust during a situation, such as: *"It felt good, was calm and comfortable",* while other responses were more detailed, for example*: "The trust was high. I am starting feeling comfortable now, I didn't even care to look right because it felt like nothing was going to happen"*. Sometimes the affective responses were of a more dual character, as when one participant expressed a high level of trust during a situation, even though the vehicle had clearly committed what was experienced as a traffic violation, stating that: "*It didn't stop at the stop sign, which I feel is strange but I still believes that it understands"*. Another participant elaborated more on the affective response by claiming that: *"On a logical level, I would like to say that I had the same trust, but the experience still felt less confident".*

In analogical responses the participants assessed the trustworthiness of the AV through comparing familiar driving styles with the driving styles experienced during the experiment. Some participants compared one of them to the driving style of people they knew. One participant explained: ”It *was more aggressive, it felt like that mate who thinks it's fun to drive fast and swing back and forth in the lane*" and another said: *”I have a son at home who is taking his license now. It was more like his driving style, maybe not the risk-taking but the jerkiness and the lane positioning”.* Yet another participant compared the driving style to the driving style of a professional driver: *“This time I think it drives like an experienced taxi driver that drives comfortably. It was a bit more stressed the first time, I prefer the calmer one”.* However, some analogical responses were of a more abstract character, comparing the driving behaviour with other familiar experiences, for example to a rollercoaster ride: “*It was fun, like the ‘grandpa cars’-ride at Liseberg”* (a rollercoaster ride located in an amusement park in Gothenburg).

During the second part of the peri-trial phase, the Trust Questionnaire resulted in a more general view on the participants’ trust for the AV whilst the ‘Trust Curve’ produced more analytic responses regarding the AVs’ trustworthiness. One participant described the driving behaviour’s effect on his/her understanding of the system’s intentions as: *”There was a distinct retardation (braking), which indicated that the system detected that there was something there, and then it braked".* Another participant explained how this understanding of the system’s intentions increased his/her perceived trustworthiness of the system: *“When the pedestrian crossed the street, I got a feeling that the car had actually seen the pedestrian and then braked softly and stopped. That is this trustworthiness that I think was lacking before”.*

Finally, the post-trial phase combination of ‘Trust Curve’, used as a mediating tool, and in-depth interviews allowed the participants to reflect on their responses. The nature of the responses during the post-trial phase were mixed in that the participants sometimes described affective responses referring to the ‘Trust Curve’ and turning them into analogical and analytic explanations.

4 DISCUSSION AND IMPLICATIONS

A number of studies have used different types of questionnaires to measure trust in technology, in general and AV in particular (e.g., Merritt et al, 2013; Haeuslschmid et al., 2017). Some, but in comparison few, have used (also) for example interviews to gather more qualitative data (e.g., Xu et al., 2014). However we argue a more comprehensive approach, including a mixed method design, to fully understand the processes of trust in an AV context.

Affective, analogical as well as analytic responses were elicited in the trial. However, different data collection methods extracted (to a certain degree) responses of different character. The methods used during the test runs extracted the majority of the affective and analogical responses, while the methods used directly after the respective test runs produced mostly analytic responses. The method used after both test runs, i.e. during the post-trial phase, elicited mixed responses and sometimes made the participants further elaborate on their earlier answers, and turning affective responses into analogical or analytic explanations.

The first data collection methods, i.e. the combination of a Likert scale and think-aloud procedure, were chosen in order to obtain direct and unfiltered answers from the participants. One explanation to the majority of affective

answers being collected during the first combination of methods (part 1) is probably that this data was collected 'in situ', i.e. during and directly after a 'critical' traffic situation. The Likert scale provided a trust measurement, indicating the participant's level of trust at the time, while the think aloud procedure allowed the participants to express their spontaneous reactions and opened up for immediate emotional responses, triggered by the situations. These affective responses would probably have been fewer if data collection had taken place after the test runs only, when the immediate emotional response may have been forgotten – affect fades faster than cognition (Norman, 2009) or is combined into an overall experience, something argued by the 'peak-and-end' rule (e.g., Fredrickson, 2000) which states that the most affectively intense event (e.g., a particular situation during a test run), together with the affect experienced at the end of a sequence (e.g., at the end of test run), affect the entirety of an experience (e.g., the entire test run). Hence, a questionnaire used at the very end of the test run may not capture the nuances of affective responses associated with different situations making it more difficult to understand participants´ trust.

The Trust Curve used directly after each test run (in part 2 of the peri-trial phase) produced mostly analytic responses. By letting the participants draw a curve from memory of their perceived level of trust and also write down the episodes that affected their trust the most, the most prominent situations were annotated. At the same time as 'peak-and-end' rule to some extent contradicts the idea of retrospective information retrieval, Karpanos et al. (2010) as well as Kujala et al. (2011) argue that such recalled memories are important since they are the most meaningful to the users. The sketching of the trust curve itself may also be of importance in recalling memories, since earlier studies have shown that when sketching chronological curves participants recalled more experiences than participants using no form of sketching (Karpanos et al., 2012). The number of analytic responses could be explained by the fact that the method allowed the participant to also reflect on the test run as a whole as well as how the individual situations shaped the overall experience. Here, the act of actually drawing the Trust Curve may also be an important characteristic of the method as such generative elements are argued as essential in order to elicit deeper knowledge about experiences (cf. Sleeswijk Visser et al., 2005).

Finally, even though the different data collection methods used in this study extracted responses of different character, it is not possible to say which particular method should be used to elicit a certain type of response. Rather the respective methods seem to complement each other in a way that enables the elicitation of richer data. This data could in turn be analysed to further explore which cognitive processes are initiated during different parts of the users' interaction with AVs (such as during the different situations) as well as in what way.

The implications are that to develop a deeper understanding, it is not sufficient to use one, single, data collection method but instead a combination of methods to elicit different aspects of the users' trust, both immediate as well as more reflective responses. Further implications are that data should not only be collected at the very end of a trial only but be complemented with data collection also during a trial, in particular in relation to events that may influence and contribute to a user's overall experience. The methods used in the study differ as to their respective affordances and constraints, some of which appropriate for immediate response (i.e. non-obstructive), others sensitising and supporting reflections. Hence, utilizing a convergent mixed method design approach could facilitate a more nuanced image of users' trust for AVs, by exploring the different cognitive processes governing trust in different situations.

REFERENCES

Charters, E. (2003). The Use of Think-aloud Methods in Qualitative Research; An Introduction to Think-aloud Methods. Brock Education Journal, 12(2), 68-82.

Creswell, J. W. & Plano Clark, V. L. (2017). Designing and Conducting Mixed Methods Research (3rd Edition ed.). Los Angeles: SAGE Publications, Inc.

Fredrickson, B. L. (2000). Extracting Meaning from Past Affective Experiences: The Importance of Peaks, Ends, and Specific Emotions. Cognition and Emotion, 14(4), 577-606.

Ghazizadeh, M., Lee, J. D. & Boyle, L. N. (2012). Extending the Technology Acceptance Model to Assess Automation. Cogn Tech Work, 14(1), 39-49.

Haeuslschmid, R., von Buelow, M. & Pfleging, B. (2017). Supporting Trust in Autonomous Driving. In Proc. of IUI 2017, March 13 - 16, 2017, Limassol, Cyprus, 319-329

Merritt, S. M., Heimbaugh, H., LaChapell J. & Lee, D. (2013). I Trust It, but I Don't Know Why: Effects of Implicit Attitudes Toward Automation on Trust in an Automated System. Human Factors 55(3), 520–534.

Jian, J.-Y., Bisantz, A. & Drury, C. (2000). Foundations for an Empirically Determined Scale of Trust in Automated Systems. International Journal of Cognitive Ergonomics, 4(1), 53-71.

Karapanos, E., Martens, J-B. & Hassenzahl, M. (2010). On the Retrospective Assessment of Users' Experiences Over Time: Memory of Actuality? CHI '10 Extended Abstracts on Human Factors in Computing Systems, 4075-4080.

Karapanos, E., Martens, J-B. & Hassenzahl, M. (2012). Reconstructing experiences with iScale. International Journal of Human-Computer Studies, 70(11), 849-865.

Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E. & Sinnelä, A. (2011). UX Curve: A Method for Evaluating Long-Term User Experience. Interacting with Computers, 23(5), 473-483.

Lee, J. & See, K. (2004). Trust in Automation; Designing for Appropriate Reliance. Human Factors, 46(1), 50-80.

Norman, D. A. (2009). THE WAY I SEE IT: Memory is More Important than Actuality. Interactions, 16(2), 24-26.

Sleeswijk Visser, F., Stappers, P. J. & van der Lugt, R. (2005). Contextmapping: experience from practice. CoDesign: International Journal of CoCreation in Design and the Arts, 1(2), 119-149.

Waytz, A., Heafner, J. & Epley, N. (2014). The Mind in the Machine; Anthropomorphism increases Trust in Autonomous Vehicles. Journal of Experimental Social Psychology , 52, 113-117.

Xu, J., Le, K., Deitermann, A. & Montague, E. (2014). How different types of users develop trust in technology: A qualitative analysis of the antecedents of active and passive user trust in a shared technology. Applied Ergonomics, 45, 1495-1503.