DESIGNING FOR CALIBRATED TRUST: EXPLORING THE CHALLENGES IN CALIBRATING TRUST BETWEEN USERS AND AUTONOMOUS VEHICLES

Valentine, David Callisto;
Smit, Iskander;
Kim, Euiyoung

Delft University of Technology

ABSTRACT

Trust is an important factor in building acceptance of autonomous vehicles within our society, but the complex nature of trust makes it challenging to design for an appropriate level of trust. This can lead to instances of mistrust and/or distrust between users and AV's. Designing for calibrated trust is a possible option to address this challenge. Existing research on designing for calibrated trust focuses on the human machine interaction (HMI), while from literature we infer that trust creation beings much before the first interaction between a user and an AV. The goal of our research is to broaden the scope of calibrated trust, by exploring the pre-use phase and understand the challenges faced in calibration of trust. Within our study 16 mobility experts were interviewed and a thematic analysis of the interviews was conducted. The analysis revealed the lack of clear communication between stakeholders, a solutionism approach towards designing and lack of transparency in design as the prominent challenges. Building on the research insights, we briefly introduce the Calibrated Trust Toolkit as our design solution, and conclude by proposing a sweet spot for achieving calibration of trust between users and autonomous vehicles.

Keywords: Communication, Design to X, Artificial intelligence, Calibrated Trust, AI Solutionism

Contact:
Valentine, David Callisto
Delft University of Technology
Industrial Design Engineering
Netherlands, The
david.cal.valentine@gmail.com

Cite this article: Valentine, D. C., Smit, I., Kim, E. (2021) 'Designing for Calibrated Trust: Exploring the Challenges in Calibrating Trust between Users and Autonomous Vehicles', in *Proceedings of the International Conference on Engineering Design (ICED21),* Gothenburg, Sweden, 16-20 August 2021. DOI:10.1017/pds.2021.114

1 INTRODUCTION

The benefits of autonomous vehicles (AV)/self-driving cars (within the paper both terms have been used interchangeably) are more speculative than verified, yet the diversity and impact of the benefits make investing in them worthwhile. The most prominent benefit of AV's is increased road safety, resulting in preventing loss of human life and reducing the economic cost incurred due to road accidents (Fagnant and Kockelman, 2015). Other benefits include improvement in fuel economy (Atiyesh, 2012), better traffic and lane management (Fagnant and Kockelman, 2015) along with allowing people in transit to put their travel time to better use (Kaelbling *et al.*, 1996; Silberg *et al.*, 2013). While the benefits of autonomous technology paint a promising picture for the future of mobility some roadblocks need to be addressed. The major one being designing an autonomous system which functions in a desired and safe manner for every situation (Campbell *et al.*, 2010), and at the same time is affordable for large-scale market adoption (Fagnant and Kockelman, 2015). In addition, challenges pertaining to the security of AV (Schoettle and Sivak, 2014), user privacy (Collingwood, 2017; Ong *et al.*, 2019), absence of clearly defined standards for testing autonomous technology (Shladover and Nowakowski, 2019), and the social consequence of the introduction of AV (Kaelbling *et al.*, 1996) still persist.

The current research focuses on the challenge pertaining to trust between users (defined as the driver in a semi-autonomous vehicle or passenger for a fully autonomous vehicle) and autonomous vehicles. More specifically exploring the challenges faced in creating an appropriate level of trust (i.e., calibration of trust) between users and autonomous vehicles. While trust is a key contributing factor in the acceptance of autonomous vehicles (Adnan *et al.*, 2018), a lack of trust (distrust) or overly trusting (mistrust) of an autonomous system can have detrimental effects (Fraedrich and Lenz, 2014). In other words, the calibration of trust is key to the acceptance of autonomous vehicles by users and safe deployment on roads. To better understand the concept of calibrated trust, we begin a literature review by exploring trust between humans and automation leading to the concept of calibrated trust that motivates our study. Subsequently, we review frameworks for calibrating trust and proceed to introduce the research gaps the paper is addressing.

1.1 Trust

Trust within literature is seen as an attitude, which guides a person's intentions and translates into behaviour within an environment and cognitive construct. Trust as an attitude is determined by beliefs and perceptions of the trustor (the person who places trust) that are formed by the availability of information of the trustee (on whom the trust is being placed) and personal experience (Ajzen and Fishbein, 1980). Lee and See (2004) describe trust as "*The attitude that an agent will help achieve an individual's goals in a situation characterised by uncertainty and vulnerability*". This definition holds for both interpersonal trust and human-automation trust. Hoff and Bashir (2015) detail the factors influencing trust between humans and automation, dividing a trust into three parts: dispositional trust (individuals' overall tendency to trust automation), situational trust (the influence of situational factors on trust), and learnt trust (trust created from interaction and past experiences). From their model, we infer that design features (in the case of an AV is the human machine interface (HMI) system) only influence the dynamic learnt trust (sub-part of learnt trust representing trust during interaction with the HMI system). In the literature, we know that trust formation starts much before the user interacts with the automation (Ekman *et al.*, 2018; Hoff and Bashir, 2015). This raises the question of how can design influence the other forms of trust i.e. dispositional trust and situational trust. It is worth pointing out that trust is not a constant and is ever-changing based on the variation in context and interaction, as highlighted by Hoffman (2017) in the statement, that the act of trusting is "*a continuous process of active exploration and evaluation of trustworthiness and reliability, within the envelope of the ever-changing work and the ever-changing system*".

1.2 Calibrated Trust

Calibrated trust is defined as the correspondence between a person's trust in automation and its capabilities (Lee and See, 2004). To understand calibrated trust it would be simpler to explore what would happen if the trust was not calibrated. In such a scenario there exist two regions: mistrust and distrust. Mistrust occurs when trust exceeds the system capabilities leading to misuse (Lee and See, 2004). Misuse or mistrust of autonomous systems has been the cause of accidents across aviation,

https://doi.org/10.1017/pds.2021.114 Published online by Cambridge University Press

maritime and automobile industries. A well-recorded incident regarding misuse of autonomous systems within an AV is the Tesla crash of 2016 (Endsley, 2019). Distrust is the opposite of mistrust, in which case users reject the capabilities of an autonomous system, the 2012 Costa Concordia cruise ship disaster is a case of distrust, as the captain of the ship deviated from the automated navigation systems in favour of manual control crashing into Le Scole reef killing 32 passengers (Lyons and Guznov, 2019). Fallon et al.(2010) proposed calibration of trust as a sensemaking process, meaning the user would need to make the decision to rely on the autonomous system based on a specific context i.e., the user must understand the autonomous function and capability followed by having the ability to predict the possible outcome of manoeuvring the autonomous system within a given context. Another way of defining calibrated trust is the active management of trustworthiness and reliability between a user and an autonomous system (Hoffman, 2017). It is important to emphasize "*active management*" because trust is not a constant nor a static property and is continuously changing, thus the calibration process is also an active and continuous process.

Operator training (in the case of AV driver training) is seen as a possible solution of building sensemaking within users (Boelhouwer et al., 2020). While this would work in the maritime, aviation and manufacturing industries within the individual mobility sector the challenge is a lot different. The minimum criteria to operate a self-driving car is owning a driver's license. In addition to this, the current technology being deployed within AV is not yet reached a satisfactory level of reliability and control (Endsley, 2019). A less than reliable autonomous system coupled with a lack of sensemaking ability in users can be seen as the prominent challenge in the calibration of trust and is aptly captured in the statement by Casner et al. (2016) *"One of the most daunting challenges will happen when we reach the crossover point where automation systems are not yet robust and reliable enough to operate without humans standing by to take over but yet are too complex for people to comprehend and intervene in a meaningful way."*

The current research into calibration of trust is mostly focused towards the interaction between the human machine interface (HMI) system and the users. One of the first approaches of calibrating trust was proposed by De Visser et al.(2014), by creating trust cues that lead to calibration of trust. These trust cues were based on the relationship between the trust evidence levels (origin, expressiveness, process, performance, and intent) and the information processing stages (perception, comprehension, project, decision, and execution). A similar framework to explore the calibration of trust was later proposed by Mirnig et al.(2016). The difference being instead of considering information processing stages, the task breakdown (operational, tactical, and strategic) were considered. Ekman et al.(2018), within their research, considered the overall journey of using an automated vehicle, from entering the vehicle to exiting it, by highlighting 13 unique interactions in the overall journey and plotting them against 11 factors that can be used to calibrate trust at the required moment. Okamura and Yamada (2020) presented a different approach towards addressing the challenge. They proposed the use of intuition and logic on one axis and machine and anthropomorphism on the other to segregate various forms of HMI.

The above-discussed research mostly focuses on the interaction between the user and the autonomous vehicle. However, we sense that trust-building starts much before the user interactions with an autonomous vehicle (Ekman *et al.*, 2018). This indicates that the process of calibration of trust starts before the user interacts with an AV. It is this gap that the current research is striving to fill, by exploring the challenges faced in the calibration of trust not just focused on the HMI and use phases but across the pre-use (begins at the product development phase and ends when the autonomous vehicle is delivered to a user) of the autonomous vehicle. Leading to the formulation of the research question as:

What are the prominent challenges faced in ensuring the calibration of trust across the pre-use phase of an autonomous vehicle/self-driving car?

The current research question has been developed to study the challenges for calibration of trust for an ownership model and does not consider challenges faced in case of shared mobility or mobility as a service (MaaS). This is because there are additional factors influencing trust in the case of shared mobility or MaaS, such as distrust towards alternate forms of mobility (Etzioni, 2019). Creating a distinction between the various ownership models will allow for a better understanding of the unique challenges each one of them would face. The research paper is part of a larger project focused on

https://doi.org/10.1017/pds.2021.114 Published online by Cambridge University Press

designing for calibrated trust in AV. The final design intervention a "Calibrated Trust Toolkit" has been briefly introduced and described in the discussion section, along with future steps.

2 METHOD

We used semi-structured interviews as an opportunity to study a diverse range of participants while at the same time providing a systematic and comprehensive understanding of the collected data (Patton, 2002). Four categories of experts were identified from the initial literature research to select as participants, these were: (experts in) autonomous vehicle technology implementation, trust between humans and automation, interaction between vulnerable road users (VRU) and AV and designing for responsible technologies. For each of these categories, separate interview guides were created based on the guidelines proposed by Patton (2002). The interview guide focused on addressing three broad questions: Current approach to designing AV's and their shortcomings, factors influencing trust in AV's and, future challenges faced by organizations in developing AV's. The sampling strategy chosen for the research was strategic and purposive (Miles *et al.*, 2014), this was achieved by creating a shortlist of possible participants based on their previously published research papers and current research areas. The second round of sampling was performed in the form of snowballing. In total 20 participants agreed to participate and 16 were interviewed (4 participants per category), the remaining 4 participants were not interviewed because of delayed response to the invitations. All interviews were conducted by Skype or Zoom and audio recorded. The average duration of the interviews was 45 minutes. Following the interviews, the audio was transcribed using Otter.ai. The second round of manual transcription was performed to check for any mistakes or errors in the transcription.

2.1 Analysis of Interviews

The interviews were analysed using Atlas.ti 8.0—a user-friendly and widespread qualitative data analysis software in academic research. The four topics were initially analysed separately and then interrelations were identified within the four separate topics. Thematic Analysis was chosen to analyse the interviews, as it allows for identifying topics that integrate into higher-order key themes (Ritchie *et al.*, 2013). The first step in the analysis was familiarizing oneself with the data. This was followed by a round of in-vivo coding and a second round of focus coding. Having completed the focus coding, themes were identified within the codes. The themes were selected based on the relevance and significance towards our primary research question. The identified themes were reviewed once to check their correlation with the initial codes (Maguire and Delahunt, 2017). The last step in the data analysis was to triangulate the data. The sequence of analysing the interviews has been represented in Figure 1.

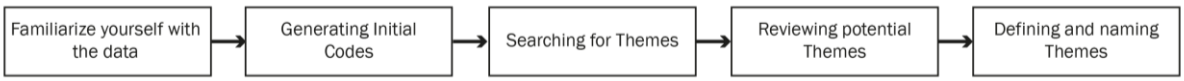

Figure 1. Representation of steps taken during the analysis of interviews

3 RESULT

This section provides an overview of the prominent challenges identified from the analysis of the interviews:

3.1 Challenges to Calibration of Trust

3.1.1 Communication

Poor communication of autonomous capabilities and limitations was seen as a major barrier in calibrating trust. Through the analysis of the interviews, three types of communication were identified that influence building and maintaining an appropriate level of trust. The overview of the relationship between the three types of communication can be seen in Figure 2.

- **Communication within the company/organization:** The first type of communication is between the various stakeholders (stakeholders consist of the product development team, manufacturing, marketing, suppliers and other functions or departments involved) present within a company. Building an autonomous vehicle is considered a complex process (Ulrich and

https://doi.org/10.1017/pds.2021.114 Published online by Cambridge University Press

Eppinger, 2012) because of the number of subsystems and stakeholders. The capabilities (conditions under which the autonomous system will function and limitations of the system) of the autonomous system must be clarified to all relevant stakeholders. This is seen to be especially challenging considering interactions between the data scientist and other stakeholders as seen by the quote:

"*data scientists have this huge lack of communication, they cannot communicate their epistemological to people that do not understand their concepts. So communication is really important, also managing this problem*" [Participant 14, RI_2]

The challenge is not only to simplify the complexity of the working and decision making of an autonomous system for other stakeholders to understand, but also to create a common direction the company wants to move in when designing the autonomous vehicle. Resulting in an overall design of an AV that provides a coherent overview of the capabilities and limitations of the autonomous vehicle for the user.

"*So you (company) need to have some kind of coherent view, you need to have the same intention from a whole company in order to design a system that can interpret in a very coherent way*" [Participant 3, Trust in AV_1]

- **Communication between company and users**: The second type is the communication between the company and users. Communication in this case generally takes place through mediums such as advertisements, online forums, or visiting the car dealership. Through the interviews, three important aspects were identified that must be communicated: *purpose of the autonomous system, limitations of the autonomous system and responsibility of the user*. The purpose of the autonomous system and limitations of the system can be seen to be quite similar but distinguishing the limiting conditions is important to prevent ambiguity in the understanding of the user. This aids in overcoming questions such as "*Should I use the autonomous system in heavy rain?*" and "*Can I use adaptive cruise control in suburban areas?*" Specifying the responsibility of the user is another aspect that is mostly overlooked but is crucial to building trust in autonomous systems. Lack/Absence of providing adequate information to the user is seen to build the habit of understanding the autonomous system by trial and error (Lin *et al.*, 2018) leading to the unwanted scenarios in which drivers were seen to push the limit of the autonomous systems capabilities. These actions not only endanger the life of the user/driver but also other road users (pedestrians, cyclists other vehicles) and tarnish the image of autonomous vehicles in the public eye (Lee and Kolodge, 2020). Possible solutions to overcome these challenges that are being investigated include, designing a human-centered perspective of explaining autonomous vehicle functions (Yang *et al.*, 2017) or providing driver training before purchasing a vehicle (Boelhouwer *et al.*, 2020).
- **Communication between the AV and users**: Calibration of trust is the active management of trustworthiness and reliability as discussed previously. Communication between the AV and user refers to the "active management", building on the previously described types of communication and consists of the design of the human machine interface (HMI) and external human machine interface (eHMI: the interfaces that communicate to other road users) system of the autonomous vehicle. It is also one of the prominent areas of research in the field of trust between users and AV's as seen in the case of De Visser et al.(2014), Ekman et al.(2018), and Okamura and Yamada (2020).

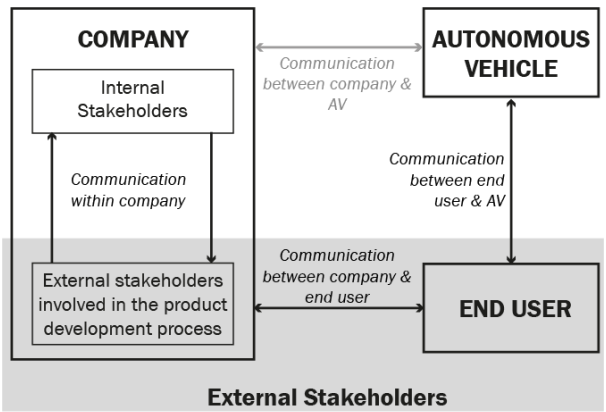

Figure 2. The figure represents the types of communication as identified from the interviews.

https://doi.org/10.1017/pds.2021.114 Published online by Cambridge University Press

3.1.2 The Solutionism Trap

The second barrier identified in designing an autonomous vehicle is the "solutionism trap". A solutionism approach, AI solutionism (Hagendorff and Wezel, 2020) or technochauvinism (Ma, 2019) is based on the assumption that technological solutions can solve social problems, by filtering the problem to a solution space within which technological solutions are viable. Such an approach can lead to negligence of human values, latent user needs and the development of unintended consequences. In essence overlooking the complexity of social problems. An alternative is a socio-technical approach to design which combines both the technological and societal aspects of the problem space when finding a solution. Rasouli and Tsotsos (2020) point out the importance of a socio-technical approach, highlighting that there is a need to move away from approaching the design of AV like a rigid dynamic problem and consider the societal effects as well. It is further pointed out in the interviews that it is up to the organisation to instil a socio-technical mind-set and approach when designing for autonomous systems, which could be aided by bringing in more industrial designers or knowledge scientists.

"*Realizing that it doesn't mean that you necessarily translate this always into some technical requirements, it's important design requirements can be socio technical in nature, and this is very important, especially in this context and socio-technical means that yes, sometimes that involves human technology interaction, but very importantly, it also involves human to human interaction that surrounds this technology, organizational processes and also sometimes the conclusion might be that you actually do not want to introduce technology into a certain situation in a certain context."* [Participant 15, RI_3]

3.1.3 Transparency of the System

Designing for transparency was listed as a possibility of providing better understandability and predictability of the autonomous system resulting in building an appropriate level of trust. Interestingly, none of the interviewees believe that there currently exists a fully transparent system. The reason suggested was based on the belief that a transparent system is created from a transparent design process. To achieve a transparent design process two criteria were highlighted that needed to be satisfied: **explainability of design choices when designing the system** and **justification of decisions made by the autonomous system**. Similar to designing with a socio-technical approach, building a transparent system also requires an attitude change within the organisation.

"So one is that the design process itself needs to be transparent, in the sense that you can see how design choices were motivated….. The second aspect of transparency that I often consider that has to do with you know, what you often see in the context of AI is referred to refer to explain ability or justification of automatically made decisions." [Participant 15, RI_3]

3.1.4 Branding and Point of Distinction

Branding plays an important role in the perception of users, especially in the case of the automobile industry (Lin *et al.*, 2018; Woisetschläger, 2016). Strong brand equity can allow automobile manufacturers to overcome the initial trepidations of users when purchasing autonomous vehicles. Two scenarios were identified within which branding influences calibration of trust:

- **Autonomous function nomenclature**: Autonomous functions within an autonomous vehicle are named to align with the brand identity. While this aids in strengthening the brand equity (the added value brands provide to a product), they do not provide the users a complete understanding of the system, as seen by the quote "*some nice words (names of autonomous functions) that look nice, like I don't know what there are, like, I think 20 different terms for adaptive cruise control*" [Participant 7, AV_1]. A similar result was also seen in the research of Abraham *et al.*(2017), where the name of the driver assistance system was seen to affect the perception of the consumers regarding the automation of the system. Tesla's choice to name their autonomous system autopilot has come under criticism for similar reasons by consumer groups and even government agencies as seen in the case of German (Nees, 2018). The interviewees emphasized the importance of aligning the nomenclature not just to the brand identity but also the capabilities of the autonomous function.
- **Brand Identity in an autonomous future**: The relationship between consumers and brands play an important role in the formation of trust (Lee and Kolodge, 2020), it is expected that a new automobiles from a manufacturer will embody the brand values of the company (e.g. Volvo is

https://doi.org/10.1017/pds.2021.114 Published online by Cambridge University Press

considered as a brand focused on safety and thus its vehicles are expected to embody this value). However, in the transition to fully autonomous vehicles, there is a risk that brands might lose their brand identity if all autonomous systems were to function in a similar manner. Ferrari for example has for the time being steered away from developing fully self-driving cars as they believe it does not fit with their brand value, instead focussing on driver assistance systems(Atiyeh, 2018). The challenge automobile manufacturer's face is integrating trust calibrating features within their products and services without compromising their brand identity.
"So they (car manufacturers) put a lot of effort on showing that they stand out in that particular way. So they don't want that they're worried that if automated vehicles are all supposed to behave within certain limitations and restrictions that every BMW also starts to drive like a grandma's car, and then the whole their selling point their USP is kind of gone" [Participant 12, VRU_3]

4 DISCUSSION

The paper aims to broaden the scope of calibrated trust by exploring the pre-use phase to understand the challenges faced in the calibration of trust. We have identified four major challenges to the calibration of trust as discussed in the results. Within this section, we continue our discussion on calibrated trust by first proposing a sweet spot for calibrating trust within the pre-use and use phase. Subsequently, we elaborate on the calibrated trust toolkit and conclude with the general findings.

4.1 Sweet Spot for Calibrating Trust

The objective of the research was to identify and develop design interventions for calibrating trust. Till now we have identified the challenges in the pre-use phase. Before we explore the final design intervention it is important to explore how the development of trust looks like for the user and what can be an ideal calibration process. Figure 3 represents the overall trust calibration process for a user across the pre-use and use phase. The trust level (blue line) is a semantic representation with mistrust and distrust as the two extremes. Ideally, we want to achieve calibration of trust before the use phase, however this is not always the case as users learn more about the AV as they interact with it in different contexts/scenarios. We also do not want to leave the calibration solely to be done in the use phase, as it puts additional design requirements for the HMI system. We thus propose a sweet spot for calibration which exists at the intersection of the pre-use and use phase (represented by the green strip). The sweet spot combines the benefits of the pre-use phase where users can accumulate information about the vehicle from various sources (e.g., reviews, advertisements, visits to dealerships and test drive) and fine tune their understanding of the system in the early stages of using the vehicle. This would additionally reduce users from extensively leaning through a trial and error approach and in certain cases prevent them from unnecessarily pushing the boundaries of the autonomous system. Ekman *et al.*, 2018 also propose a similar concept labelling it as a learning stage, further solidifying the importance of combining the pre-use and use phase in achieving calibration of trust.

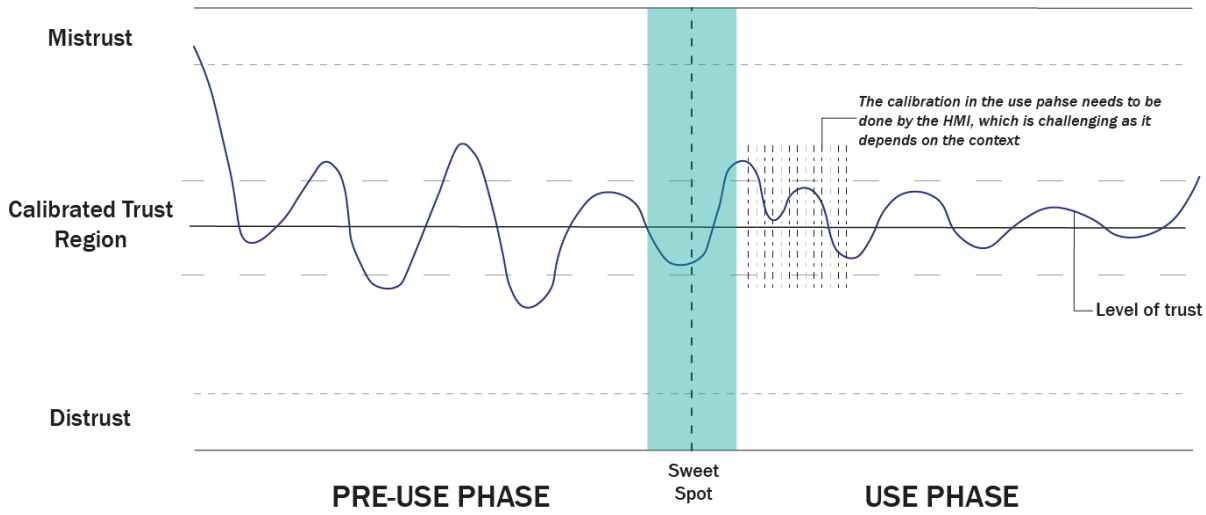

Figure 3. A semantic representation of the sweet spot for achieving calibration of trust

https://doi.org/10.1017/pds.2021.114 Published online by Cambridge University Press

4.2 Calibrated Trust Toolkit

A toolkit was selected as the design intervention for its relevance and impact at an organizational level of human experience (Friedman and Hendry, 2019; Rao et al., 2020). The Calibrated Trust Toolkit (Figure 4) has been developed to be employed by the product development team, and is based on the two prominent challenges identified i.e. communication and building a socio-technical approach to designing an AV. The toolkit consists of four parts. The first is a sensitizing session which builds an understanding of the factors influencing trust and the concept of calibrated trust to the product development team. The second part is an Autonomous Function Visualization Canvas which provides a graphical and human-centered representation of the autonomous systems for simplified understanding of the system for stakeholders. The third part is the User Decision Matrix, that facilitates the exploration of the influence of contextual factors on drivers and their interaction with the AV. The last part of the toolkit is the Trust Enhancing Communication, which is a framework to support communicating autonomous capabilities to stakeholders. The toolkit has undergone an initial round of testing, but further testing and validation is yet to be conducted.

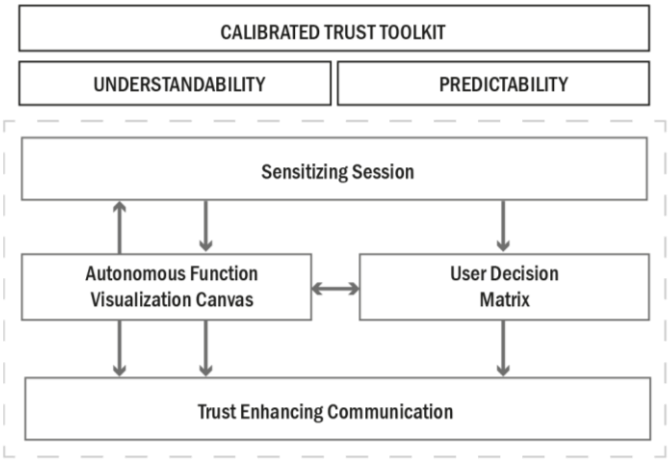

Figure 4. Overview of the calibrated trust toolkit

4.3 General Findings

The interviews provided a holistic understanding of the challenges faced in the calibration of trust. The identified challenge aligns with the overarching barrier that was highlighted at the beginning of the research paper pertaining to building trust between a less than reliable technology and users that do not completely understand the working of autonomous systems. The results can be summarised by considering two prominent clusters:

- **Communication**: The bedrock of calibration of trust is information exchange, this not only includes technical working of the AV but also other information such as product name/nomenclature. Companies need to thus be careful in the manner in which they represent their autonomous technology to users. At the same time find a universal nomenclature to explain the technical working of autonomous systems.
- **Approach to Design**: How companies approach the designing of autonomous systems directly influences the manner in which users will interact with the automation. A socio-technical approach towards design was considered as an important starting point along with focusing on building a transparent design approach. Building transparency is also connected to communication as it allows for explaining the decision making of the system to the user.

5 CONCLUSION

Trust is an important part in creating acceptance of autonomous systems within our society. Through this research we have tried to underline some challenges that are present in the calibration of trust by exploring the pre-use phase of an autonomous vehicle. The result shows that there is a need to focus on communicating the autonomous capabilities and limitations of AV to the various stakeholders within the pre-use and use phase. Also, facilitate a shift towards a socio-technical approach to designing autonomous systems. The influence of branding on trust while evident in hindsight

https://doi.org/10.1017/pds.2021.114 Published online by Cambridge University Press

highlighted the complexity of designing for trust and the various factors that influence its calibration. The research by no means covers the depth and complexity of trust in automation. We thus, encourage further research into exploring the challenges and possible design interventions that can be employed in calibrating trust. For inspiration we have highlighted three avenues for future research extending from our research:

- **Influence of ownership models**: How do ownership models influence the trust formation process and how would calibration of trust occur in these scenarios?
- **Interaction between use and pre-use phase**: How does the interaction with an HMI system during the use phase influence the pre-use phase? Or how do the activities performed within the pre-use phase influence the interaction between the user and HMI within the use phase?
- **Communicating capabilities of AV**: How might we communicate the capabilities and limitations of AV to stakeholders (stakeholders consist of the product development team, manufacturing, marketing, suppliers and other functions or departments involved)? And how can we measure the extent of the success in their communication?

REFERENCES

Abraham, H., Seppelt, B., Mehler, B. and Reimer, B. (2017), "What's in a Name", Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, ACM, New York, NY, USA, pp. 226–234.

Adnan, N., Md Nordin, S., bin Bahruddin, M.A. and Ali, M. (2018), "How trust can drive forward the user acceptance to the technology? In-vehicle technology for autonomous vehicle", Transportation Research Part A: Policy and Practice, Vol. 118, pp. 819–836.

Ajzen, I. and Fishbein, M. (1980), "Understanding attitudes and predicting social behaviour", Englewood Cliffs.

Atiyeh, C. (2018), "Don't Expect a Self-Driving Ferrari", Car and Drive, available at: https://www.caranddriver.com/news/a23362999/ferrari-no-to-autonomous-cars/ (accessed 7 October 2020).

Boelhouwer, A., van den Beukel, A.P., van der Voort, M.C., Verwey, W.B. and Martens, M.H. (2020), "Supporting drivers of partially automated cars through an adaptive digital in-car tutor", Information (Switzerland), available at:https://doi.org/10.3390/INFO11040185.

Campbell, M., Egerstedt, M., How, J.P. and Murray, R.M. (2010), "Autonomous driving in urban environments: Approaches, lessons and challenges", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, available at:https://doi.org/10.1098/rsta. 2010.0110.

Casner, S.M., Hutchins, E.L. and Norman, D. (2016), "The challenges of partially automated driving", Communications of the ACM, available at:https://doi.org/10.1145/2830565.

Collingwood, L. (2017), "Privacy implications and liability issues of autonomous vehicles", Information and Communications Technology Law, available at:https://doi.org/10.1080/13600834.2017.1269871.

Ekman, F., Johansson, M. and Sochor, J. (2017), "Creating Appropriate Trust for Autonomous Vehicle Systems: A Framework for Human-Machine Interaction Design", 95th Annual Meeting of the Transportation Research Board.

Ekman, F., Johansson, M. and Sochor, J. (2018), "Creating Appropriate Trust in Automated Vehicle Systems: A Framework for HMI Design", IEEE Transactions on Human-Machine Systems, Institute of Electrical and Electronics Engineers Inc., Vol. 48 No. 1, pp. 95–101.

Endsley, M.R. (2019), "Situation awareness in future autonomous vehicles: Beware of the unexpected", Advances in Intelligent Systems and Computing, available at:https://doi.org/10.1007/978-3-319-96071-5_32.

Etzioni, A. (2019), "Cyber Trust", Journal of Business Ethics, Springer Netherlands, Vol. 156 No. 1, pp. 1–13.

Fagnant, D.J. and Kockelman, K. (2015), "Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations", Transportation Research Part A: Policy and Practice, Elsevier Ltd, Vol. 77, pp. 167–181.

Fallon, C.K., Murphy, A.K.G., Zimmerman, L. and Mueller, S.T. (2010), "The calibration of trust in an automated system: A sensemaking process", 2010 International Symposium on Collaborative Technologies and Systems, CTS 2010, No. October 2016, pp. 390–395.

Fraedrich, E. and Lenz, B. (2014), "Automated driving: Individual and societal aspects", Transportation Research Record, available at:https://doi.org/10.3141/2416-08.

Friedman, B. and Hendry, D.G. (2019), Value Sensitive Design, Value Sensitive Design, The MIT Press, available at:https://doi.org/10.7551/mitpress/7585.001.0001.

Hagendorff, T. and Wezel, K. (2020), "15 challenges for AI: or what AI (currently) can't do", AI & SOCIETY, Vol. 35 No. 2, pp. 355–365.

Hoff, K.A. and Bashir, M. (2015), "Trust in automation: Integrating empirical evidence on factors that influence trust", Human Factors, Vol. 57 No. 3, pp. 407–434.

https://doi.org/10.1017/pds.2021.114 Published online by Cambridge University Press

Hoffman, R.R. (2017), “A taxonomy of emergent trusting in the human-machine relationship”, Cognitive Systems Engineering: The Future for a Changing World, pp. 137–164.
Kaelbling, L.P., Littman, M.L. and Moore, A.W. (1996), “Reinforcement Learning: A Survey”, Transportation Research Board Annual Meeting, available at: http://arxiv.org/abs/cs/9605103.
Lee, J.D. and Kolodge, K. (2020), “Exploring Trust in Self-Driving Vehicles Through Text Analysis”, Human Factors: The Journal of the Human Factors and Ergonomics Society, Vol. 62 No. 2, pp. 260–277.
Lee, J.D. and See, K.A. (2004), “Trust in automation: Designing for appropriate reliance”, Human Factors, Vol. 46 No. 1, pp. 50–80.
Lin, R., Ma, L. and Zhang, W. (2018), “An interview study exploring Tesla drivers’ behavioural adaptation”, Applied Ergonomics, Elsevier, Vol. 72 No. December, pp. 37–47.
Lyons, J.B. and Guznov, S.Y. (2019), “Individual differences in human–machine trust: A multi-study look at the perfect automation schema”, Theoretical Issues in Ergonomics Science, available at:https://doi.org/10.1080/1463922X.2018.1491071.
Ma, K. (2019), “Artificial unintelligence: How computers misunderstand the world”, The Information Society, pp. 1–3.
Maguire, M. and Delahunt, B. (2017), “Doing a Thematic Analysis: A Practical, Step-by-Step Guide for Learning and Teaching Scholars. All Ireland Journal of Teaching and Learning in Higher Education”, All Ireland Journal of Teaching and Learning in Higher Education n (AISHE-J).
Miles, M.B., Huberman, A.M. and Saldana, J. (2014), Qualitative Data Analysis: A Methods Sourcebook. Third Edition, The SAGE Handbook of Applied Social Research Methods.
Mirnig, A.G., Wintersberger, P., Sutter, C. and Ziegler, J. (2016), “A framework for analyzing and calibrating trust in automated vehicles”, AutomotiveUI 2016 - 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Adjunct Proceedings, pp. 33–38.
Nees, M.A. (2018), “Drivers’ perceptions of functionality implied by terms used to describe automation in vehicles”, Proceedings of the Human Factors and Ergonomics Society, Vol. 3, pp. 1893–1897.
Okamura, K. and Yamada, S. (2020), “Adaptive trust calibration for human-AI collaboration”, PLoS ONE, Vol. 15 No. 2, pp. 1–20.
Ong, A., Troncoso, J., Yeung, A., Kim, E. and Agogino, A.M. (2019), “Towards Flexible Ridesharing Experiences: Human-Centered Design of Segmented Shared Spaces”, Communications in Computer and Information Science, available at:https://doi.org/10.1007/978-3-030-23525-3_50.
Patton, M.Q. (2002), “Qualitative research and evaluation methods. Thousand Oaks”, Cal.: Sage Publications.
Rasouli, A. and Tsotsos, J.K. (2020), “Autonomous vehicles that interact with pedestrians: A survey of theory and practice”, IEEE Transactions on Intelligent Transportation Systems, IEEE, Vol. 21 No. 3, pp. 900–918.
Ritchie, J., Lewis, J., Nicholls, C.M. and Ormston, R. (2013), “The foundations of qualitative research. Qualitative research practice: A guide for social science students and researchers”, Qualitative Research Practice: A Guide for Social Science Students and Researchers.
Schoettle, B. and Sivak, M. (2014), “A survey of public opinion about connected vehicles in the U.S., the U.K., and Australia”, 2014 International Conference on Connected Vehicles and Expo, ICCVE 2014 - Proceedings, available at:https://doi.org/10.1109/ICCVE.2014.7297637.
Shladover, S.E. and Nowakowski, C. (2019), “Regulatory challenges for road vehicle automation: Lessons from the California experience”, Transportation Research Part A: Policy and Practice, available at:https://doi.org/10.1016/j.tra.2017.10.006.
Silberg, G., Manassa, M., Everhart, K., Subramanian, D., Corley, M., Fraser, H., Sinha, V., et al. (2013), “Self-Driving Cars: Are we Ready?”, Kpmg Llp.
Ulrich, K.T. and Eppinger, S.D. (2012), Product Design and Development: Fifth Edition, McGraw-Hill.
De Visser, E.J., Cohen, M., Freedy, A. and Parasuraman, R. (2014), “A design methodology for trust cue calibration in cognitive agents”, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 8525 LNCS No. PART 1, pp. 251–262.
Woisetschläger, D.M. (2016), “Consumer perceptions of automated driving technologies: An examination of use cases and branding strategies”, Autonomous Driving: Technical, Legal and Social Aspects, available at:https://doi.org/10.1007/978-3-662-48847-8_32.
Yang, J.H., Han, J. and Park, J.M. (2017), “Toward defining driving automation from a human-centered perspective”, Conference on Human Factors in Computing Systems - Proceedings, Vol. Part F1276, pp. 2248–2254.

https://doi.org/10.1017/pds.2021.114 Published online by Cambridge University Press