Calibrated Trust – A Prerequisite for a Better Automation Usage and User Experience in Highly Automated Driving Context

Kaiyu Huang
Technical University Munich
Institute of Ergonomics
Boltzmannstr. 15
85747 Garching
Germany
kaiyu.max.huang@gmail.com

ABSTRACT
How people trust a system determines how he would use the automation. In highly automated driving (HAD) context, a driver uses an automation only if he trusts it. However, overtrust, blind-trust, distrust may cause abuse, misuse, disuse respectively. Hence, a proper trust which matches the system capability, is the prerequisite for the defined use scenario and desired user experience (UX). Ways to improve trust system calibration like a real experience, system transparency, mechanism understanding, criterion setting were proposed in this essay. Following this, further considerations and notices about optimizing trust, such as low fidelity information delivery, humanization, demographical and cultural differences were discussed. Finally, we had an outlook on trust engineering as a future profession.

KEYWORDS
Trust; automation; user experience; automated driving;

1 INTRODUCTION: HAD and UX
Highly automated driving is developing rapidly in the recent years. In the 2020s, 2-5% of the sold cars can drive autonomously. The market of this new technology will finally be saturated in the 2060s, according to the study (Litman, 2014). Car manufacturers and solution providers have been investing a lot into HAD research for a greater market share in this huge competition of future mobility. User experience is playing an important role apart from the technology.

Before talking about UX in HAD, the definition and facts of UX should be clarified. The term User Experience refers to a person’s perceptions and responses that result from the use or anticipated use of a product, system or service according to the ISO definition. However, the multidisciplinary nature of UX has led to several definitions of and perspectives on UX, each approaching the concept from a different viewpoint (Roto, Vermeeren, Väänänen-Vainio-Mattila, & Law, 2011). As claimed by Hassenzahl, Diefenbach, & Göritz (2010), UX is rather fun-oriented. It is derived from the hedonic quality of a product. Usability on the other hand is more goal-oriented. It shows the pragmatic quality of the product. Usability is the precondition of a good UX. Nevertheless, the presence of a satisfying usability doesn’t necessarily mean a good UX.

To the question, what causes a good UX, Hassenzahl et al. (2010) believe the UX quality is positively correlated with the fulfillment of human needs. Furthermore, Hassenzahl and colleagues found out the seven most salient psychological needs of human in a technology product context (after their priority): relatedness, stimulation, competence, popularity, security, meaning, autonomy. It reveals the sensation seeking nature of human as a social animal. Interestingly, security is seen as a less important need. This can be explained by understanding it as a “deficiency need”, i.e., the absence of security causes a decrease of UX, while the presence doesn’t necessarily mean a big UX improvement.

2 TRUST IN AUTOMATION
2.1 Why is Trust So Important in HAD?
The importance of human as a decision maker has never been so highlighted before the automation era. As the automation paradox describes: "the more efficient the automated system is, the more essential the human contribution that is needed to run the Automation System. Humans are less involved in heavily automated systems, but their involvement becomes more critical" (Kaufman, 2010). Several research have indicated that the decision
whether user uses an automated function is dependent upon the trust in automation (Muir, 2007; Parasuraman & Riley, 2016). How people trust a system determines how he would use the automation. Improper trust leads to improper use of the product, hence an improper UX. More specifically, overtrust, blind-trust, distrust may cause abuse, misuse, disuse respectively. No matter how sophisticated the automation is, it can be rejected if human as the decision maker doesn’t trust it. In the HAD context, disuse of the automation resulted from a low trust, means the potential UX benefits from the liberation of driving task and from non-driving related tasks are lost. If the user is forced to use to automation which he doesn’t trust, a lack of security, as well as a high monitive workload will happen, thus probably a poor UX. Consequently, trust as a prerequisite for UX is standing on a higher level than UX influencing factors.

However, the operator’s use of automatic controllers depends upon more than trust alone (LEE & MORAY, 1992). For example, a lack of perceived usefulness could be one reason of disuse (Trübswetter & Bengler, 2013). Besides, automation tends to be followed by a human operator in an emergency scenario, even with a prior automation failure (Verberne, Ham, & Midden, 2012).

### 2.2 Understand of Trust

#### 2.2.1 What is Trust?

Trust was defined according to Lee & See (2004) as the attitude that an agent will help achieve an individual’s goals in a situation characterized by uncertainty and vulnerability. It is rather an emotional attitude than a psychological need. This emotional attitude plays a bigger role in decision making, especially in an imperfect automation with unanticipated circumstances where cognitive resources are not available to support a calculated rational choice. A common belief was that the trust between human and machine is different with in-human trust. Interestingly, Jian, Bisantz, & Drury (2000) have found out the same nature between human-human trust and human-machine trust. They also stated that trust is a one-dimensional variable. Conversely, Spain, Bustamante, & Bliss (2008) have revealed that trust is actually a multi-dimensional construct with two comprising distinct, yet related, factors: trust and distrust. That is to say, a low distrust doesn’t necessarily mean a high level of trust. These two findings are important for defining approaches in trust engineering. On the other hand, there is still some differences between human-human trust and human-machine trust, e.g., the trust between humans is actually a social exchange relationship with the interplay of offering and expecting, while the automation system doesn’t expect anything from the human.

#### 2.2.2 Measurement of Trust

One direct way to measure automation trust quantitatively is to use the self-report measurements. Questionnaires like Automation Trust Scale (Jian et al., 2000), Single-item Automation Trust Ratings (Brown & Galster, 2004) are used. Besides, research has discovered the dynamic nature of trust (Yang, Unhelkar, Li, & Shah, 2017), which shows trust as a variable state, should be measured continuously. A trust survey at the end of an experiment would be inadequate. Thus, trust of entirety would be better quantified by the average measure of “area under the trust curve” than the traditional post-experiment trust measure. Trust can also be quantified by user behavior objectively. One of the most important approaches in HAD is to measure the gaze behavior. The standard deviation of horizontal gaze distribution (Gold, Körber, Hohenberger, Lechner, & Bengler, 2015), glance for monitoring per second during the non-driving-related tasks (Hergeth, 2016) are proved to be valid methods. Reaction time could also be considered as a trust quantification, assuming that decisions are made quickly if an operator trust or distrust the system strongly (Itoh, Abe, & Tanaka, 1999).

### 3 TRUST SYSTEM CALIBRATION

#### 3.1 What Kind of Trust Do We Want?

Trust works differently in different SAE levels. In SAE level 2, human is still responsible for the dynamic driving task (manual driving mode) and motive tasks (automated driving mode). In this less reliable system, where Tesla Autopilot1 stays, overtrust may lead to accidents due to insufficient monitoring. While a distrust causes a high cognitive strain of monitoring for the human driver. In SAE level 3, human is relieved from monitive tasks. Yet human driver as a fallback level, has to still take over the control, as long as the automation system cannot handle the situation. In this level, trust can still impact UX dramatically. In a low trust level, drivers are not able to fully enjoy the benefits of NDRT and automation. Again, overtrust may still lower the take-over quality and causes an accident. In SAE level 4 and 5, trust would play a less role in UX during the automated driving, since human driver is relieved completely from the driving related tasks. However, trust is a precondition for making the decision to use the automation. User experience won’t even exist, if users don’t trust the system, thus don’t use the system.
An interesting video showed a person’s first time experience with the Tesla Autopilot (CrownFleek, n.d.). Amazement, fun, scare were shown during the driving. According to Hassenzahl’s need fulfillment model, the insufficient security gain as a “deficiency need” should be critical to UX. However, we can hardly say that was a pretty bad UX. The reason of this can be an extreme level of stimulation fulfillment. Nevertheless, even though the first experience causes a high momentary UX, but for the long run, the cumulative UX must be strongly harmed due to the low security gain or even accidents. Moreover, the driving experience also depends on personality traits. In this video, if the test driver had a “failure-avoiding motivated” character, instead of an “achievement motivated” personality (what he actually has), the UX would be completely the opposite. The age factor cannot be ignored as well.

Incidents of overtrust and distrust are generally found in some research (Robinette, Li, Allen, & Howard, 2016). Humans treat machine in a sense like treating human beings. It is exactly the mechanism how automation trust comes from. It is inevitable for human as a living being to have incorrect understandings of automation, even if they are designed by us.

The mismatch of trust and system capabilities is one of the reasons for the incorrect use of automation. A high level of trust doesn’t definitely indicate a good UX. Overtrust, for example, can lead to a good UX temporarily. But once the system fails, the damaged UX could be even lower than a distrustful use. While a low trust also doesn’t mean a bad UX either. A proper trust, based on the fact of system capabilities, is what we want. The calibration between trust and automation reliability (trustworthiness) is crucially important for a correct use of automation, as a UX prerequisites.

### 3.2 Ways to Build Trust System Calibration

A correct calibration of trust and system capability makes correct expectations. Once we expect automation correctly, it is more likely to have a good UX in the right design frame. Ways to improve the trust calibration include:

#### 3.2.1 Real Experience

Without a real experience of using the system, misunderstanding or prejudice happens easily. We can hardly say the automation trust is correct without experiencing it. A simulator study has its limitations for studying trust in HAD. The differences between them can be: reality of the simulation, foreknowledge of the experimental environment, participativeness in the simulation etc. But a simulator study might be the best solution at the moment to reveal the trust and behavior in real HAD.

#### 3.2.2 Mechanism Understanding

Automation trust is like human trust, which is based on understanding. If the user understands the underlying mechanism how the automation works and how decisions are made, the trust correctness will be improved. However, in an HAD context, systems are usually programmed like a black box, which makes it challenging.

#### 3.2.3 System Transparency

Moreover, knowing the current system status can help improve trust-system calibration. The system transparency can be concluded into information transparency and goal transparency (Verberne et al., 2012). Information transparency refers to the conditions of current status. It might be a speed or battery indicator, or traffic conditions around. On the other hand, the goal transparency deals with what the system is doing and what is the intention next. System transparency reduces the user uncertainty.

Nevertheless, the system uncertainty e.g., a hesitation of predicting the traffic, should be presented to the user timely. Being honest as a crucial source of trust in human, also applies to human-automation trust. A study also showed that, TOR lowers the trust only temporarily (Hergeth, Lorenz, & Krems, 2015). In a long run, trust increases again due to understanding and prediction of the system. The presentation of system uncertainty improves the human-automation cooperation when failure happens. Automation failures do not decrease trust if they are known in advance (Beggiato & Krems, 2013). In other words, informing about automation failures before they occurred leads to a better acceptance of the automation as being imperfect. Thus, the presentation of automation uncertainty cushions the reliability lost (Beller, Heesen, & Vollrath, 2013).

Another considerable approach to present system transparency is visualization. Showing visualized system interpretation of the current traffic situation can improve automation trust (Häuslschmid, Bülow, & Pflüging, 2017). This can be explained by, visualization builds better situational awareness and mode awareness for the driver, hence a trust improvement. The info-graphic design and interaction design in HAD are accordingly becoming important.
However, by improving system transparency, information overload should be avoided. Overloaded information leads to a cognitive stress to process. Besides, the quantity of presented information also depends on the time period of product experience. Detailed information at the learn phase can be an UX contribution, but it can also be annoying for an experienced user (Verberne et al., 2012). Consequently, only necessary information should be presented at the right use phase.

3.2.4 Correct Criterion Setting. Criterion setting in alarm notification is critical. For safety reasons, criteria are often set very low and lead to many false alarms. Oppositely, a high criterion causes few false alarms but many misses. An example here is the smoke alarm. Some people may turn it off due to the automation “cry wolf” effect, where human operators begin to reject an automated system due to repeated false alarms. “Cry wolf” effect makes operators question the system or even to abandon the automation. We should note that, fostering trust is a lot more difficult than destroying it. On the other hand, automation transparency may mitigate the “cry wolf” effect (Yang et al., 2017), but also brings about extra cognitive processing.

3.2.5 Considering Trust and Distrust in Calibration Separately. Since trust is a two-dimensional construct comprising with trust and distrust, the calibrations for trust and distrust should also be considered separately. For example, understanding the underlying system mechanism tends to be a method to foster trust, while information and goal transparency can be understood as a tool to correct distrust.

3.2.6 Training. Training helps build proper expectations and acceptance of automation. Incorrect expectations lead to disuse or abuse of the system. Simultaneously, training improves the system understanding and helps user perceive more usefulness of the system. The will of use is hence increased. On the other hand, the trust lost resulted from incorrect use is reduced after training.

3.3 Other Considerations for Optimizing Trust

When we talk about improving trust, it is necessary to distinguish whether it means improving the correctness of trust capability calibration, or improving calibrated trust itself. We should also think about are we trying to lower overtrust or raising distrust. Here are some considerations about improving calibrated trust.

3.3.1 Pragmatic Quality. As we conclude, trust should not exceed the actual system reliability, and should be calibrated with system capability. We assume, the usability or rather pragmatic quality of an automation, is the foundation of trust. This can be translated into system performance and the occurrence of faults (LEE & MORAY, 1992). A higher automation reliability results in a higher value of trust. Thus, improving trust without improving the system capability and objective reliability would be a nonsense.

3.3.2 Low Fidelity Information Delivery. The study shows, binary system feedback shows a better trust gain compared with a likelihood feedback (Yang et al., 2017). It confirms that “high-likelihood alerts engender a greater increase to momentary trust upon automation success, as well as a greater decline in momentary trust upon automation failure”. Accordingly, ambiguity can be advantageous for building a steady trust and UX. It makes calibration process more efficient. A reason of this can be a lower cognitive workload. Important is, a low fidelity information delivery is not on the opposite side of system transparency. It claims rather a qualitative information transparency (e.g. graphic) instead of a quantitative one (e.g. numeric). Thus, information ambiguity might be an alternative strategy to improve user experience (Jung, Sirkim, Gür, & Steinert, 2015) and speed up trust calibration process.

3.3.3 Humanization. As discovered (Jian et al., 2000), human-human trust and human-automation trust share the same nature. By humanizing the automation, especially the interaction, trust can be improved. The J.A.R.V.I.S. in Iron Man, or rather the operating system in movie “Her” can be very good examples of building trust by humanization. There is no doubt that it is easier for human as emotional animal, to trust another “human”. Specifically, the humanization can be presented by human voice, anthropomorphic visualization, or even personality. This can be a way to compensate the difficulty of understanding the underlying mechanism of a nowadays programmed smart-system. However, it is crucially important for system designers not to design the humanization misleadingly. The humanization should match the actual system reliability.

3.3.4 Demographical, Gender, and Cultural Difference. The effects of age, gender, and culture in trust building are generally found. Those differences are important for international car manufacturers to adapt and promote their products. The study found, elderly drivers are more positive to HAD than young drivers (Gold et al., 2015). Furthermore, male drivers have a higher trust than females. Hergeth et al. (2015) further found that, Chinese driver has a higher mistrust level towards HAD than German drivers. These findings can be interesting for policy and law makers as well.

3.4 Trust Engineering, A Profession in the Future?

As trust being so important, there will probably be specialists who work on optimizing human technology trust in the future. The task of these trust engineers is to give solutions to make human technology trust calibration more accurate, efficient and user-friendly. Their work domain is not restricted in automation, e.g. HAD, but in a bigger picture of technology, for instance, online-banking and e-commerce. Their academic backgrounds can be design, psychology, human factors engineering etc.
Specifically, trustworthy user interface and info-graphics should be defined by UI and UX designers. While industrial designers deliver trust by form, material, and other design languages subconsciously. On the other side, psychologists and human factors engineers build trust models and refine their theories by experiment. Moreover, as for a perspective on training, educators and media workers can also contribute to this profession by delivering correct knowledge of technology to the public.

REFERENCES


Hergeth, S., Lorenz, L., & Krems, J. F. (2015). Effects of take-over requests and cultural background on automation trust in highly automated driving, ... On Human Factors ....


