Supporting Trust in Autonomous Driving

Renate Haeuslschmid[1] **Max von Buelow**[2] **Bastian Pfleging**[1] **Andreas Butz**[1]

[1] {renate.haeuslschmid, bastian.pfleging, butz}@ifi.lmu.de
University of Munich (LMU)
Munich, Germany

[2]max.vonbuelow@online.de
IAV GmbH
Munich, Germany

ABSTRACT

Autonomous cars will likely hit the market soon, but trust into such a technology is one of the big discussion points in the public debate. Drivers who have always been in complete control of their car are expected to willingly hand over control and blindly trust a technology that could kill them.

We argue that trust in autonomous driving can be increased by means of a driver interface that visualizes the car's interpretation of the current situation and its corresponding actions. To verify this, we compared different visualizations in a user study, overlaid to a driving scene: (1) a chauffeur avatar, (2) a world in miniature, and (3) a display of the car's indicators as the baseline. The world in miniature visualization increased trust the most. The human-like chauffeur avatar can also increase trust, however, we did not find a significant difference between the chauffeur and the baseline.

ACM Classification Keywords

H.5.2 Information interfaces and presentation (e.g., HCI): User Interfaces

Author Keywords

Automated driving; trust; trustworthiness; avatar; anthropomorphism.

INTRODUCTION

Automated driving is a highly discussed topic in the broader public. One future goal for the car industry is to provide vehicles that adapt "to the drivers' needs in terms of a pleasurable and authentic driving experience" [32]. Different surveys, however, showed diverse attitudes towards this technology. A recent survey showed that on the one hand, people are fascinated by but hesitate to trust autonomous driving: 43% of the participants stated that they are afraid of driving in an autonomous car [23]. A survey by Schoettle & Sivak [35] revealed that 22% of their participants could not imagine riding in a fully automated car. Similarly, Kyriakidis et al. [19] found that 65% of their participants were worried about the reliability of autonomous cars.

IUI 2017, March 13 - 16, 2017, Limassol, Cyprus

ACM 978-1-4503-4348-0/17/03$15.00
DOI: http://dx.doi.org/10.1145/3025171.3025198

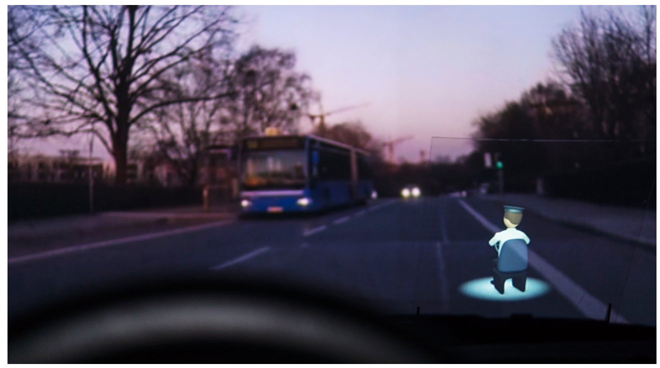

Figure 1. Our participants experienced a real driving scene seated in a real car in order to achieve a realistic feeling of autonomous driving. The intelligent autopilot visualizations – in this picture the chauffeur avatar – reacted to defined situations in the driving scene in order to convey competence and thereby foster trust.

Up to now, such figures and surveys should be considered carefully, as most participants have not yet experienced any automated driving functions, which makes a true judgment very difficult. However, already today, we see that at least some specific automated cars already show the technological feasibility, also highlighting their benefits: For instance, Google lists the statistics of their self-driving cars[1] as follows: Between 2009 and May 2015 and 10,000 autonomously driven miles per week there were only 12 and only non-severe accidents which were not caused by the self-driving car; in 2016 there was the first crash caused by their autonomous car [30]. Even though no complete statistics exist (many light accidents remain unreported), this accident rate is expected to be considerably lower than the one of a normal driver [2].

We see that trust is an important factor in the judgment of autonomous systems and influences the using behavior [12, 18, 27]. Hence a supportive user interface is essential – especially in the transition phase towards automated driving where the "driver" needs to give up control in favor of an unknown feature. The results of our study supports this phase and helps manufacturers to establish trust of the novice user in the automated vehicle.

In particular, we compare three different visualizations presented on a head-up display – a chauffeur avatar, a world in miniature, and the car's indicators – in order to understand

[1]http://www.google.com/selfdrivingcar/files/reports/report-0515.pdf, last access of URLs: 2017-01-08
[2]https://backchannel.com/the-view-from-the-front-seat-of-the-google-self-driving-car-46fc9f3e6088

whether they can positively influence the driver's feeling of trust. We decided to present those on a head-up display since it enables faster attention switches between the driving scene and the supportive visualization.

BACKGROUND & RELATED WORK

Understanding the interaction between humans and automated cars is still challenging since most drivers have not yet had the chance to experience an automated car. In the remainder of this section, we give an overview of related work on the interplay between user and automation. For driving automation, we mainly address the highest level of driving automation (i.e., SAE level 5, [34]) where the car takes over full control.

Drivers, Passengers & Automation

Driving simulator setups allow to investigate specific aspects of the interaction between the driver and the (simulated and automated) vehicle. However, driving simulations often lack external validity (in particular behavioral validity) since participants might behave differently in the real world [3]. One way to circumvent this issue is to perform costly real-world driving experiments. However, given that automated cars are not yet available, workarounds need to be found. One example is the real-road autonomous driving simulator, which represents a wizard-of-oz car [1, 2]: In this case an instructed driver steers the car manually (in a right-handed car for countries driving on the right side of the road) but is hidden from the actual participant by a separating wall. The wizard driver is expected to simulate the automated driving. While this approach enables a certain degree of automation simulation, the wizard's driving style obviously influences how the autonomous ride is perceived.

Besides investigations of how drivers (want to) interact with vehicle automation [5, 8, 29] and especially how control is shared between car and driver (e.g., [11, 36]), lately various researchers investigate issues related to driver's and passenger's trust in vehicle automation [19, 20, 23, 31, 37]. One specific aspect is the investigation of the relationship between anthropomorphic perception of the automation functions and the driver's trust in the system [37]. We discuss this aspect in the following sections.

As explored by Yusof et al. [40], another interesting aspect of autonomous driving is the driving style of the autonomous car (e.g., acceleration). They assume that the users' preferences and tasks (e.g., performing certain non-driving-related activities) affect the favored driving style. Overall, they identified a general trend to accept defensive driving, despite it might contradict the driver's own driving style.

Anthropomorphism

As outlined by Epley et al. [37] anthropomorphism is the attribution of human-like features to, e.g., objects or animals [10]. Humanization in this regard can either be related to the appearance or the nature of the object. This does not only include the attribution of human character traits but also emotional acting and complex intellectual capacity. Epley et al. [10] furthermore outline that an anthropomorphic perception enables humans to understand objects by attributing pre-existing knowledge and structures. This positively influences comprehension and helps to establish an emotional relation.

Several experiments already revealed a relation between anthropomorphism and trust [6, 13, 26, 33]. Waytz et al. [38] studied anthropomorphism in order to increase trust in autonomous driving by means of human-like voice output in a dynamic driving simulation. They found that anthropomorphism leads to an increased trust in fully automated driving. However, the power of this effect, especially compared to other channels (e.g, visual representations as we use them in our experiment) remains unclear. Therefore, we compare the effects of different visualizations for automated driving: an anthropomorphic *chauffeur avatar*, a computer-vision approach called *world in miniature*, and a basic visualization of the car's turn intentions by its *indicators*.

Motivated by prior work which indicated that people do not like to interact with simplistic agents due to a mismatch between a simplistic deployment and high user expectations, Parise et al. [28] investigated how users cooperate with speech-enabled software agents. They showed that the interaction with agents represented as dogs was perceived as more pleasant but at the same time, less cooperation took place compared to human agents. The latter might be caused by the perceived agent competence.

Trust: Psychological Models & Theories

Literature provides a set of definitions of the abstract concept of trust, including its meaning as a personality trait, belief, social structure, or behavioral intention. Often, the term is used to describe interpersonal relations [33]. Several psychological theories have been developed to model trust. We used this as a starting point when identifying opportunities to increase trust through different visualizations in our experiment.

Mayer et al. [24] found that the perceived trustworthiness is based on the situationally perceived integrity, benevolence, and ability. They identified personality traits and prior experiences with the trusted person/object as moderating factors. Hoff & Bashir [16] described three layers of reliability in human-automation trust, namely dispositional trust, situational trust, an learned trust, along with assumptions about the trusted object/person. This includes perceived competence, benevolence, integrity, and predictability. Based on their findings, they built a model for conceptualizing the variability of trust with the goal to support the construction of trust.

McKnight & Chervany [14] divided trust into preferences towards trust, situational influences, and trust that is created by using a system. In contrast to other literature, they also propose concrete steps how to increase the trust in a certain automation. This includes anthropomorphism, the increase of usability, a polite communication, transparent system actions, along with the chance to intervene in the system's activities.

All these models highlight the importance of the situation in which trust is needed along with the characteristics of the truster and the trustee. One important aspect for the construction of trust are feedback loops that result from continuously using a system. Mayer et al. [24] distinguish in their cyclic trust model between trust and the actual performance of a trust-motivating action. The result of this action influences trustworthiness and, thus, closes the loop. This means that the perceived trust is adjusted whenever this cycle is followed and that a basis for the next decision is created when a trust-motivated action is performed.

Hoff & Bashir [16] use a similar cyclic model. In their model, dynamically learned trust and the use of a system are influenced by constant factors (e.g., system properties), the current situation, and flexible factors (e.g., the perceived performance of an automation). The existence and influence of a feedback loop depends on the actual system. For systems with only little automation, users will decide for each action of the system whether they should trust its advice. This is also caused by the fact that in these situations consequences of system actions can often be easily rated, which enables adaptation of the interaction with the system. The more automation a system provides, the less a user can judge and monitor the quality of the automation. This follows from the complexity of the automated tasks and the reduced interaction with the system.

Trust in Autonomous Systems

For interpersonal trust, the propensity to trust is often seen as a separate personality trait. Hoff & Bashir assume that a similar effect can also be observed for trust into (automated) systems [16]. However, questionnaires about interpersonal trust are highly related to the relationship between two persons, which makes it difficult to transfer this to the trust in automated systems. With this regard, Merit et al. [25] found a relation between a person's general attitude to machines and how they use it in different situations. Their questionnaire is used to measure the basic trust in machines.

According to Hoff & Bashir [16] trust is influenced by three components, (1) the person who trusts, (2) the system this person is supposed to trust, and (3) the situation. The first component (i.e., the *person*), is characterized by the propensity to trust, which is influenced by different factors (e.g., gender, age, opinions, character traits). In our experiment, we observe these factors but do not systematically compare them through a systematic selection of participants.

For the second component (the system to be trusted) it is important that trust is influenced by the subjectively and situationally perceived integrity, benevolence, and ability of the system. By continuously using the system, the user collects experiences about the system's functionality and performance and adjust her or his trust towards this system. This highlights that trust is a dynamic construct.

The situation (3rd component) is considered independent of the system. It describes the overall circumstances in which the trust relationship should be established. The influence of the situation is particularly characterized by the underlying risk of this situation. In the context of our work, the situation is defined by the properties of a fully automated ride and our test setup, which is designed to feel as realistic as possible.

The trust-influencing features of systems as explained by Mayer et al. [24] and McKnight & Chervany [14] can be roughly divided into two groups. Competence, ability, and predictability are closely correlated to the performance of an automation. In contrast, integrity and benevolence are purely emotional aspects.

Trust in Automated Cars

Existing display concepts in partly automated cars visualize the state of the system and, thus, let the driver monitor the automated components. Regarding trust-influencing factors, these concepts improve the ability of the system and support predictability, which in turn generates trust. Display concepts which also address emotional trust factors could positively influence trust. Hoff & Bashir [16] also provide a set of recommendations on how to design automated systems in order to increase the user's trust. This includes using anthropomorphism, increasing usability, polite communications by the system, transparent system behavior, as well as the opportunity for the user to intervene.

Lee et al. [20] conducted an ethnographic experiment by observing participants riding in a prototype of a level 2-automated car on real roads. In their evaluation, they identified nine factors of distrust related to the dimensions of trust as defined by Lee and Moray [21]: system performance, the automation process (algorithms), and purpose. Based on their findings, they propose ideas on how to reduce distrust in automation. They assume that the driver's anxiety about unpredictable situations may be reduced once they observe that the car accurately manages various driving situations. In addition, they propose the use of an agent for a positive impact on trust, e.g., by increasing the emotional connection between driver and agent. They suggest the agent to take different appearances and also support or entertain the driver beyond driving activities.

Hergeth et al. [15] followed another approach to understand trust in automation and user behavior. They investigated the relationship between gaze behavior and automation trust. They found a negative relationship between the self-reported trust in automation and the driver's frequency of monitoring the automation.

In order to understand the passengers' mental conditions, emotional states, and opinions, Wintersberger, Riener, & Frison [39] investigated how the choice of the operator (male or female driver or driving automation) affects the passengers. By conducting a simulator study they found that the choice of the driver only has little effect on emotional states and mental conditions. As a consequence, they conclude that passengers are inclined to already accept automated cars.

AUTOPILOT VISUALIZATIONS

We designed three visualizations for autonomous driving: a chauffeur avatar, a world in miniature and the car's indicator. The car's indicator only communicates the car's intentions to turn. The world in miniature is a computer vision style visualization and presents the car's perception of the surroundings, its interpretation and its actions in a clean and competent way. The anthropomorphic visualization, the chauffeur avatar, reacts to the same events as the world in miniature but is more human-like and potentially associated with more feelings.

Situation	World i. Miniature	Chauffeur Avatar
turn	indicator symbol	glance & turns indicator on
close objects	sensor symbol	looks at object
traffic light	traffic light symbol	looks at traffic light
danger	warning sign	glance & exclamation mark

Table 1. We designed two of the intelligent visualizations (world in miniature and chauffeur avatar) to represent equivalent levels of intelligence. Both visualizations react to the same events.

The number of actions or symbols respectively as a reaction to the environment influences the perceived competence of the system. We decided to only visualize important, safety-relevant situations and actions in order to avoid clutter and overwhelming the user. The reactions of both visualizations are depicted in table 1.

Chauffeur Avatar

Anthropomorphic visualizations can present complex systems in an easy and friendly way [38]. However, if the simulation is inappropriate, people can react with dislike and reject the system. A high competence seems to be more important than symphathy [28]. Users do not enjoy interacting with too simple avatars. While for an online shop, realistic avatars might be more likeable, appropriate and trustable [22], such highly realistic avatars can create high expectations which, if not satisifed, can severely harm acceptance. In general, avatars that are similar to the own person, e.g., in look, behavior and gender are preferred [4, 22]. Friendly behavior such as smiling can further enhance the overall perception of and the trust in the avatar [4].

Based on these findings, we developed our own comic style avatar. Corresponding to our passenger situation and the expected participants, we decided on a male chauffeur avatar of mid-age and with European taint and professional clothing. Our chauffeur was animated with frequent friendly behavior: He waves for greeting and turns around to the driver and smiles when the driving situations allows. We enhanced our avatar in six cycles of informal discussions with potential users to make it look friendly and competent.

The 3D model of our avatar was created in Cinema 4D[3]. We developed a toolkit of short animations and motions and connected these to sequences according to the driving videos. The chauffeur's motions were animated by means of the motion capture sequences of Adobe Mixamo[4] which led to a very human and natural look.

World in Miniature

The display of the surroundings as a world in miniature is already a common visualization for advanced driver assistant systems in high-class cars. The *world in miniature* communicates its understanding of the surroundings as well as its actions by an animated visualization of an abstract 3D model of the real world. It is based on the display of the road and the own car and can further include, for example, other cars, speed limits, lane boundaries, road exits, traffic signs and lights. By representing the world correctly and acting appropriately, the system proves its competence and performance, which are important factors of trust.

We based our world in miniature on the one used by Tesla[5] and extended it with visualizations specific for inner-city driving such as intersections. As for the avatar, we created several basic 3D models and animations by means of Cinema 4D and connected these pieces according to the driving videos. As for the cars, we used 3D models from DMI[6] and cgtrader[7]; we matched the 3D referent to the own car with the model of the car used in the study setup. Example visualizations of the world in miniature are depicted in the figures 1 and 2.

Car Indicators (Baseline)

As a baseline, we used a basic visualization of the car's indicators instead of no visualization. The indicators are a standard and due to safety-aspects also a mandatory function for manual and in future most probably also for autonomous driving. They do not communicate any intelligence or understanding of the situation but give feedback about the car's intentions so that turns do not happen unexpectedly which could affect the trust negatively. As the other visualizations, the indicator arrows are displayed on the HUD (see figure 2).

EXPERIMENT

Prior to the study reported below, we performed a brief pilot study with 6 participants in order to test the transfer of the level of trust from one test phase into the other and also to evaluate the study design and the test setup. Below, we only report the final experiment.

Research Questions & Hypotheses

The study was designed to evaluate the initial trust in autonomously driving cars and to investigate weather a supporting intelligent visualization could increase trust. Based on the related work, we assumed that an intelligent visualization that seems to understand the situation increases the drivers' trust due to the competence it communicates. Further, we hypothesized that a human-like visualization, such as the *chauffeur avatar*, would increase trust more than a computer vision style visualization, such as the *world in miniature*, because in a driving scenario humans normally trust other humans, e.g., as the co-driver of a friend or in a taxi. We also assumed that trust is directly linked to the perceived safety.

Participants

We recruited 30 participants (12 female) by means of online social media and e-mail. We did not pre-select participants according to specific criteria other than a valid driver's license. Our participants had a mean age of 26 years (SD=5.9) and were compensated with vouchers.

[3]https://www.maxon.net/en-us/products/cinema-4d/overview/
[4]https://www.mixamo.com/
[5]www.tesla.com/de_DE/autopilot
[6]www.dmi-3d.net
[7]www.cgtrader.com

Figure 2. We designed three visualizations for autonomous driving cars. *Left*: A world in miniature, *Middle:* A chauffeur avatar, *Right:* Basic car indicators as the baseline. The world in miniature and the chauffeur avatar are intelligent autopilot visualizations that interpret the current driving situation and react accordingly. The car indicator visualization only visualizes the basic turn intentions and does not show any intelligence. This corresponds to what drivers currently see in cars. We ensured that the indicators light up and sound for the same duration in all autopilot variants.

We measured the participants' attitudes and behaviors by several questionnaires (see section Questionnaires). In general, our participants seemed to be safe drivers. However, many of them would cross an intersection although the traffic light is about to turn red and exceed the speed limit for up to 15 km/h but not 30 km/h. They would not do risky overtaking maneuvers or race for fun or out of boredom.

Overall, our participants trust autonomous systems until they fail or show deficits in competence. They are also confident that autonomous systems have a high competence. In addition, all of our participants are thrilled about and like to use novel technologies. Comparing the attitude towards autonomous driving of our participants with the results of Kyriakidis et al. [19], we found that our participants were slightly more fascinated about fully autonomous driving. In both studies, autonomous and manual driving are rated as equally comfortable. We therefore think that our set of participants is representative but, probably due to the technical background of many of them, slightly positively biased about autonomous driving.

Study Design

We designed a within-subjects experiment with 3 x 3 conditions: We counterbalanced three driving videos and three autopilot visualizations (chauffeur, world in miniature, indicators) throughout the study, resulting in three test phases. We defined six groups that represent the presentation orders of the autopilot visualization (at fixed order of the driving videos) and randomly assigned five participants to each group.

The driving videos were recorded by means of a GoPro HERO 4 (1280 x 960 px resolution) which was placed inside the car below the rear view mirror. We drove the same track several times during different times of day and weather conditions to obtain similar but not identical footage. We then cut three driving videos with 7 min each out of two different recordings: The first part shows urban two-lane roads with little to medium traffic density. It contains many maneuvers such as stopping at a traffic light and turns and situations such as crossing pedestrians and overtaking cyclists. The second part shows high-density traffic on an urban multi-lane road with a lot of lane change traffic and many stops at traffic lights. It also contains one unpredictable event that requires fast action such as a suddenly stopping lead car.

For each driving video, we then designed corresponding simulations for each autopilot visualization. The visualizations *chauffeur* and *world in miniature* present their intelligence by their understanding of and reaction to objects, events and situations. To ensure comparability, both visualizations react to the same events within each driving video. The baseline visualization only communicates the car's turn intentions by means of the *indicators*; this visualization does not convey intelligence or understanding of the situation.

In order to create a high feeling of realism, we used the videos of real driving along with a real car test setup. Participants even started the automated drive themselves by pressing a button on a smartphone next to the steering wheel, which also aimed to increase the feeling of control and the interactivity. In order to keep participants involved in the situation and prevent distraction, the experimenter frequently asked the participant to judge the current feeling of trust on a scale from 1 to 10 (low to high trust). Overall, the study lasted about 75 min.

Questionnaires

Since we investigated subjective feelings, we based our evaluation on questionnaires. Prior to the test, we used *introductory questionnaires* to gather demographic data as well as to get insights into the participants' attitudes and tendencies to trust people or systems. After each test phase, participants had to fill out the *intermediate questionnaires* to report their feeling of trust in autonomous driving supported by one particular autopilot visualization. After the last test phase and its intermediate questionnaire, the participants had to fill out a *closing questionnaire* which compares the three visualizations directly.

All information was collected by means of a tablet and online forms. Below are the detailed contents of the questionnaires:

Introductory Questionnaires

Demographic Data

Basic information such as age and gender, driving experience, use of transportation and assistance systems

Driving Behavior

9 risky driving behaviors that evaluate the participants' tendency to take a risk or not; extracted from the 'young adult driving questionnaire' from Donovan & Jessor [9] which is a specialization of the SSSV-TAS [7]

Trust in Autonomous Systems

tendency to trust autonomous systems; questionnaire from Merrit et al. [25]

Attitude towards Novel Technologies

9 statements with 5 point Likert scales that evaluate a person's attitude towards and usage of novel technologies

Attitude towards Autonomous Driving

7 statements with 5 point Likert scales that evaluate the participants' attitude towards autonomous driving; extracted from Kyriakidis et al.'s questionnaire [19]

Intermediate Questionnaires

Confidence & Safety

10 statements with 5 point Likert scales evaluating the participants' confidence in the system and the perceived safety

Trust

12 statements with 5 point Likert scales about the feeling of trust in the autonomous system and the autopilot; trust questionnaire from Jian et al. [17] which was used in many studies about trust [16]

User Experience

26 opposing adjectives with 7 point Likert scales which belong to the criteria attractiveness, perspicuity, efficiency, dependability, stimulation and novelty; standard user experience questionnaire (UEQ)[8]

Closing Questionnaire

Comparison of the Autopilot Visualizations

Each visualization was rated on a 7 point Likert scale regarding safety, understandability, reliability, trust, aesthetics and likability for each autopilot visualization. Participants had to select their favorite visualization.

Procedure

The experimenter welcomed the participants and instructed them to sit down in the car. For a realistic feeling they were asked to adjust the seat position and to close the seat belt. Then, the experimenter introduced the participants to the study procedure and handed them the tablet to fill out the introductory questionnaires. The experimenter asked people to think aloud during all drives and to report their current trust in the autopilot at the experimenter's prompt. Then, the experimenter set up the first drive. The autopilot greeted the participants who could then start the drive by pressing the start button on the smartphone interface next to the steering wheel. After the drive, participants had to fill out the intermediate questionnaires. This procedure was repeated for each of the two remaining visualizations. After the last drive, participants filled out the closing questionnaire, shared personal thoughts and gave feedback. The experimenter then released the participants from the study.

[8]www.ueq-online.org

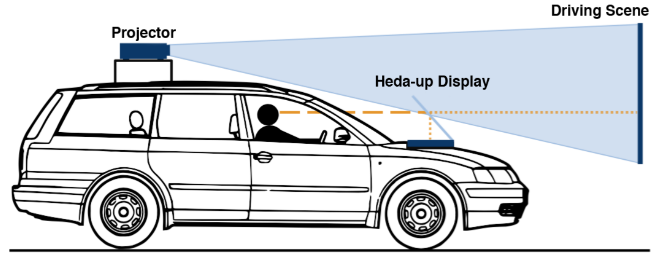

Figure 3. The projector was placed on the roof of a Volkswagen Passat so that the projected image covered the entire field of regard through the windshield.

Apparatus

Most models of trust mention the perception of the situation in which one has to trust a person or a machine as an important factor. Hoff and Bashir [16] describe the situational trust and Mayer et al. [24] mention the perceived risk as important factors of their model. Hence, the test situation and setup need to be as realistic as possible in order to enable a valid and meaningful evaluation of trust.

We placed a Volkswagen Passat in a garage and played back videos of inner-city driving. The driving videos were projected on the wall in front of the car so that the driving scene is of realistic size and the entire field of regard through the windshield is covered. Further, we played back the sound of the driving videos from a speaker placed inside the car.

The head-up display was simulated by a Microsoft Surface Pro 3 (12 inch, 2160 x 1440 px resolution) that was placed below an acrylic glass pane. We selected a very thin glass pane of 3 mm thickness to avoid double image effects. The glass pane was mounted at an angle of 45° by a small wooden retainer and transparent fishing lines so that the HUD itself was as inconspicuous as possible. We placed this construction on the car's hood so that the reflected image was approximately 1.5 m in front of the driver. Our HUD simulation provided a bright and sharp image (see figure 1 and 4).

RESULTS

Trust

We measured the participants' trust in an autopilot directly after the test drive by the questionnaire of Jian et al. [17] on a scale from 0 to 4. The miniature world visualization was rated best with a value of mean=2.9, followed by the avatar visualization (mean=2.7) and the indicator visualization (mean=2.4). We performed a Friedman test with an adjusted α-level=0.017 as well as post-hoc pairwise comparisons (Wilcoxon test, α-level=0.05) and did not find a statistically significant difference for the trust level measured directly after each test phase.

Figure 4. The head-up display was designed to not interfere with the participants' view of the road scene and shows a clear and bright image.

Then, we evaluated the trust in the three autopilots after all visualizations were tested. We found a significant main effect (p<.001) for trust x visualization. The post-hoc tests revealed a significant difference between baseline and miniature world (p<.001), between miniature world and avatar (p=.024) and also between avatar and baseline (p=.039). We think that the differences in the results of the two measurements can be attributed to the time of measurement. The second measurement of trust was performed after all visualizations had been experienced and hence participants did reflect and compare the single visualizations.

Safety

As trust, the safety of the autopilots was evaluated after each test phase as well as after all three test phases. The participants rated the single autopilots with a safety of mean=3.0 for the *world in miniature*, mean=2.3 for the chauffeur, and mean=2.4 for the indicator visualization. The Friedman test showed a significant interaction effect between the visualizations and the perceived safety (p=0.014). A Wilcoxon post-hoc analysis showed a significant difference between *world in miniature* and *baseline* (p=.002) and a nearly significant difference between baseline and avatar (p=.06).

After the participants experienced all three visualizations, they had to rate the safety of all variants again. The results confirmed the significant main effect of visualization x safety. Wilcoxon pairwise comparisons showed a significant difference between *world in miniature* and *chauffeur avatar* (p=.024), between *world in miniature* and baseline (p<.001) and also between *chauffeur* and *baseline* (p=.039).

User Experience

Our user experience measurements further support the order of the three visualizations (see figure 5): The *world in miniature* visualization received the highest ratings for all UEQ criteria, followed by the *chauffeur avatar*. The *baseline* visualization received the lowest user experience ratings. We compared the user experience of the three visualizations with the UEQ benchmark, which is based on UEQ results of 163 studies. The *world in miniature* and the *avatar* received comparably good values for perspicuity and novelty, average values for attractiveness, efficiency and dependability and low values for stimulation.

Direct Comparison

Participants had to answer six questions about safety, understandability, reliability, trust, aesthetics and likability for each autopilot visualization. Throughout all questions, participants rated the *world in miniature* visualization with the highest values, followed by the *chauffeur avatar*. As the final question, we asked participants which autopilot visualization they would like to use in their autonomous car. 60% of our participants (n=18) stated to prefer the *world in miniature*, 20% (n=6) would like to use the *chauffeur avatar* and 13% (n=4) chose the basic indicator visualization. Two participants decided against any visualization.

We analyzed the the results for these questions by Friedman tests and found significant differences for likability (p=0.012), aesthetics (p=0.003), and understandability (p=0.003). Wilcoxon post-hoc tests showed a significant difference between baseline and miniature world (p<.001 / p<.001 / p<.001) and between miniature world and avatar (p=.02 / p=.03 / p=.004) for all three criteria.

Individual Feedback

We asked our participants to think aloud and give continuous feedback about the autopilots. Some participants stated that it really felt like they were driving autonomously and that they completely forgot that they were only in a garage. The combination of the stationary car setup and the driving video did barely lead to a feeling of motion sickness; only one participant reported to feel sick. Many participants said that they missed a navigation system that shows the travel route. Some participants stated that the driving behavior of the autopilots was not inspiring confidence, surprisingly, because of the defensive behavior. They interpreted the defensive driving behavior as uncertainty and stated that a more competent autopilot would drive faster in some situations. In contrast, few participants thought that the autopilots drove aggressively. Regarding the *chauffeur avatar*, many participants expected different behavior; the opinions ranged from passive and phlegmatic to active and hectic. Also, the entire setup was described as too playful to be safe. Interestingly, there seem to be two groups of users: Participants either stated that they felt safe if there was a visualization that shows a clear understanding of the situation or that they felt safer if there was no visualization at all.

DISCUSSION & LIMITATIONS

Prior studies pointed at a lack of trust in autonomous cars [19, 23]. We assumed that an intelligent visualization of the car's understanding and interpretation of the situation as well as its corresponding actions could increase trust by creating transparency and communicationg competence.

We developed three autopilot visualizations – a world in miniature, a chauffeur avatar, and the car indicators – and animated them according to the driving scenes. These animations were synchronized with the driving videos and played back on a head-up display so that the image floated above the car's hood. We decided to use a HUD because it simplified the observation of both, the driving scene and the autopilot visualization, but think that any other in-car display would be suitable as well.

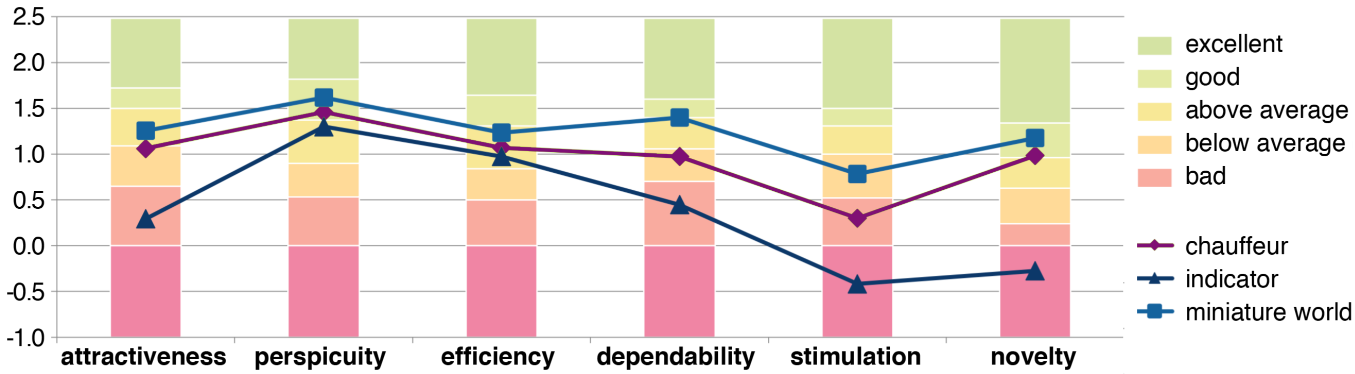

Figure 5. The world in miniature autopilot visualization achieved overall good user experience values, closely followed by the chauffeur avatar. The baseline visualization of the car indicators received the lowest UEQ values. All variants are easy to understand but not stimulating enough.

Autonomous driving was simulated by playing back videos of real driving in front of a real car our participants were sitting in to create a highly realistic setting and feelings. Though some of our participants explicitly stated that they felt as if driving in a real autonomous car, they were aware that there is no real jeopardy and hence, the perceived safety of and the trust in the simulated autopilot might be influenced by the setting and differ from a real autonomous car.

The risk of an accident and the feeling of loss of control certainly becomes more real for the passenger of an autonomously driving car. Since none of our participants had ever experienced this so far, it is very hard for them to estimate how much they would actually trust the single autopilot visualizations. We can imagine that at first use, people desire an autopilot that is likeable and with which they can interact; possibly to be distracted from their fears. This autopilot does not necessarily have to be a human-like visualization.

In general, our participants showed slightly higher (initial) trust in autonomous driving compared to the survey by Kyriakidis et al. [19]. One reason for this might be that our group of participants was rather young and did not include elderly drivers. Another might be that many of them had technical backgrounds. As expected, both intelligent visualizations increased trust compared to the basic presentation of the car's indicators. The results show a clear and consistent order of the three visualizations regarding the efficiency to increase trust and perceived safety as well as user experience: The *world in miniature* autopilot visualization has the strongest effects, followed by the *chauffeur avatar*. The *car indicator* visualization is the least effective variant.

We do not think that the graphical quality of the avatar is the reason for the lower trust in this visualization due to our iterative process to improve the look of the avatar and the precise animation of reactions to the outside world. This assumption is supported by the very similar values for user experience. We assume that the miniature world conveys a stronger feeling of competence and is more suitable to the technical system than a human-like visualization. The avatar was expected to create a stronger feeling of trust in the sense of comfort and well-being instead of competence. It seems that competence is a more important factor, potentially due to the safety-critical situation, but further research is needed to verify this hypothesis.

Surprisingly, the opinions about the autopilots' driving behavior varied considerably: While some participants found it aggressive, many others judged it as defensive. Both, a defensive but also an aggressive autopilot counteract trust and perceived safety. Also, participants did not agree on the need of a visualization: They either stated that they felt safer if there is a visualization that shows a clear understanding of the situation or that they felt safer if there is no visualization at all. These statements show, that the needs of the future users vary considerably and suggest that a one-fits-all visualization might not be sufficient. A visualization that adapts to the user's own driving behavior could be a good solution, especially if the autopilot is represented by an avatar.

CONCLUSIONS

Autonomous driving is about to become a reality. Although people are fascinated by this novel technology, they are also skeptical and worry about its actual safety. We think that a visualization of the autopilot can help people to build trust in the autonomously driving car. We gathered insights into the propensity of trust and attitudes of our participants and compared three different autopilot visualizations and their influence on trust, perceived safety, and user experience. One visualization represents the chauffeur of the car: We developed a *chauffeur avatar* which was thought to increase trust because of its human-like look and behavior. Further, we developed a visualization that represents the world in miniature; a visualization that is already used for driving assistance systems and that was assumed to communicate high competence. As the baseline visualization, we used the car's indicators, which provide a minimal feedback about the car's intentions to turn.

As expected, the anthropomorphic visualization increased the trust in the autonomous car. However, the world in miniature had considerably stronger effects on the participants' trust. It also fostered the strongest feeling of safety as well as the best user experience. Since the results of our study clearly favor the world in miniature and most participants stated that they wanted such a visualization in their autonomous car, we recommend to further investigate this visualization. Future research should study to which extent the visualization has to represent and react to the current driving scene. Also, a large augmented reality visualization on the windshield might be an interesting advancement. In addition, a larger and more diverse group of participants should be involved since, for example, older people are expected to have a different attitude towards this novel and potentially dangerous technology.

ACKNOWLEDGMENTS

This work was partially funded by and executed in cooperation with IAV GmbH.

REFERENCES

1. Sonia Baltodano, Srinath Sibi, Nikolas Martelaro, Nikhil Gowda, and Wendy Ju. 2015a. The RRADS Platform: A Real Road Autonomous Driving Simulator. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '15)*. ACM, New York, NY, USA, 281–288. DOI: http://dx.doi.org/10.1145/2799250.2799288

2. Sonia Baltodano, Srinath Sibi, Nikolas Martelaro, Nikhil Gowda, and Wendy Ju. 2015b. RRADS: Real Road Autonomous Driving Simulation. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts (HRI'15 Extended Abstracts)*. ACM, New York, NY, USA, 283–283. DOI: http://dx.doi.org/10.1145/2701973.2702099

3. Gary E. Burnett. 2009. On-the-Move and in Your Car: An Overview of HCI Issues for In-Car Computing. *International Journal of Mobile Human Computer Interaction (IJMHCI)* 1, 1 (2009), 60–78. DOI: http://dx.doi.org/10.4018/jmhci.2009010104

4. Andrew J. Cowell and Kay M. Stanney. 2003. Embodiment and Interaction Guidelines for Designing Credible, Trustworthy Embodied Conversational Agents. In *Proceedings of Intelligent Virtual Agents: 4th International Workshop (IVA 2003)*, Vol. 2792. Springer Berlin Heidelberg, 301–309. DOI: http://dx.doi.org/10.1007/978-3-540-39396-2_50

5. Rita Cyganski, Eva Fraedrich, and Barbara Lenz. 2015. Travel-Time Valuation for Automated Driving: A Use-Case-Driven Study. In *Annual Meeting of the Transportation Research Board*. 11–15.

6. Ewart J. de Visser, Frank Krueger, Patrick McKnight, Steven Scheid, Melissa Smith, Stephanie Chalk, and Raja Parasuraman. 2012. The World is not Enough: Trust in Cognitive Agents. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56, 1 (2012), 263–267. DOI: http://dx.doi.org/10.1177/1071181312561062

7. Patricia Delhomme, Nadine Chaurand, and Françoise Paran. 2012. Personality predictors of speeding in young drivers: Anger vs. sensation seeking. *Transportation Research Part F: Traffic Psychology and Behaviour* 15, 6 (Nov. 2012), 654–666. DOI: http://dx.doi.org/10.1016/j.trf.2012.06.006

8. Murat Dikmen and Catherine M. Burns. 2016. Autonomous Driving in the Real World: Experiences with Tesla Autopilot and Summon. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 16)*. ACM, New York, NY, USA, 225–228. DOI: http://dx.doi.org/10.1145/3003715.3005465

9. J. Donovan and R. Jessor. 1992. Young adult driving questionnaire. *University of Colorado Institute of Behavioral Science, Boulder* (1992).

10. Nicholas Epley, Adam Waytz, and John T Cacioppo. 2007. On seeing human: a three-factor theory of anthropomorphism. *Psychological Review* 114, 4 (Oct 2007), 864–886. DOI: http://dx.doi.org/10.1037/0033-295X.114.4.864

11. Yannick Forster, Frederik Naujoks, and Alexandra Neukum. 2016. Your Turn or My Turn?: Design of a Human-Machine Interface for Conditional Automation. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 16)*. ACM, New York, NY, USA, 253–260. DOI: http://dx.doi.org/10.1145/3003715.3005463

12. David Garcia, Christine Kreutzer, Karla Badillo-Urquiola, and Mustapha Mouloua. 2015. Measuring Trust of Autonomous Vehicles: A Development and Validation Study. In *Proceedings of HCI International 2015, Part II - Posters' Extended Abstracts (HCI International 2015)*, Constantine Stephanidis (Ed.). Springer International Publishing, Cham, 610–615. DOI: http://dx.doi.org/10.1007/978-3-319-21383-5_102

13. Li Gong. 2008. How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior* 24, 4 (July 2008), 1494–1509. DOI: http://dx.doi.org/10.1016/j.chb.2007.05.007

14. D. Harrison McKnight and Norman L. Chervany. 2001. *Trust and Distrust Definitions: One Bite at a Time*. Springer Berlin Heidelberg, Berlin, Heidelberg, 27–54. DOI: http://dx.doi.org/10.1007/3-540-45547-7_3

15. Sebastian Hergeth, Lutz Lorenz, Roman Vilimek, and Josef F. Krems. 2016. Keep Your Scanners Peeled. *Human Factors* 58, 3 (2016), 509–519. DOI: http://dx.doi.org/10.1177/0018720815625744

16. Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57, 3 (2015), 407–434. DOI: http://dx.doi.org/10.1177/0018720814547570

17. Jiun-Yin Jian, Ann M. Bisantz, and Colin G. Drury. 2000. Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics* 4, 1 (March 2000), 53–71. DOI: http://dx.doi.org/10.1207/S15327566IJCE0401_04

18. John D. Lee and Katrina A. See. 2004. Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46, 1 (2004), 50–80. http://hfs.sagepub.com/content/46/1/50.short.

19. M. Kyriakidis, R. Happee, and J.C.F. de Winter. 2015. Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour* 32 (July 2015), 127–140. DOI: http://dx.doi.org/10.1016/j.trf.2015.04.014

20. Jiin Lee, Naeun Kim, Chaerin Imm, Beomjun Kim, Kyongsu Yi, and Jinwoo Kim. 2016. A Question of Trust: An Ethnographic Study of Automated Cars on Real Roads. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 16)*. ACM, New York, NY, USA, 201–208. DOI: http://dx.doi.org/10.1145/3003715.3005405

21. John Lee and Neville Moray. 1992. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics* 35, 10 (1992), 1243–1270. DOI: http://dx.doi.org/10.1080/00140139208967392

22. J. T. Luo, Peter McGoldrick, Susan Beatty, and Kathleen A. Keeling. 2006. On-screen characters: their design and influence on consumer trust. *Journal of Services Marketing* 20, 2 (2006), 112–124. DOI: http://dx.doi.org/10.1108/08876040610657048

23. Matthias Schmidt. 2016. Unfälle mit selbstfahrenden Autos: Für Viele ist der Hersteller verantwortlich. https://yougov.de/news/2016/07/19/unfalle-mit-selbstfahrenden-autos-fur-viele-ist-de/. (July 2016).

24. Roger C. Mayer, James H. Davis, and F. David Schoorman. 1995. An Integrative Model Of Organizational Trust. *Academy of Management Review* 20, 3 (1995), 709–734. DOI: http://dx.doi.org/10.5465/AMR.1995.9508080335

25. Stephanie M. Merritt, Heather Heimbaugh, Jennifer LaChapell, and Deborah Lee. 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 (June 2013), 520–534. DOI: http://dx.doi.org/10.1177/0018720812465081

26. Richard Pak, Nicole Fink, Margaux Price, Brock Bass, and Lindsay Sturre. 2012. Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. *Ergonomics* 55, 9 (Sept. 2012), 1059–1072. DOI: http://dx.doi.org/10.1080/00140139.2012.691554

27. Raja Parasuraman and Victor Riley. 1997. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39, 2 (1997), 230–253. DOI: http://dx.doi.org/10.1518/001872097778543886

28. Salvatore Parise, Sara Kiesler, Lee Sproull, and Keith Waters. 1996. My Partner is a Real Dog: Cooperation with Social Agents. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work (CSCW '96)*. ACM, New York, NY, USA, 399–408. DOI: http://dx.doi.org/10.1145/240080.240351

29. Bastian Pfleging, Maurice Rang, and Nora Broy. 2016. Investigating User Needs for Non-Driving-Related Activities During Automated Driving. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia (MUM '16)*. ACM, New York, NY, USA, 91–99. DOI: http://dx.doi.org/10.1145/3012709.3012735

30. Associated Press. 2016. Google self-driving car caught on video colliding with bus. *The Guardian* (March 2016). https://www.theguardian.com/technology/2016/mar/09/google-self-driving-car-crash-video-accident-bus.

31. Andreas Riener. 2014. Who Cares About Trust, Grade of Traveling & Quality of User Experience in a World of Autonomous Cars?. In *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '14)*. ACM, New York, NY, USA, 1–3. DOI: http://dx.doi.org/10.1145/2667239.2667311

32. Christina Rödel, Susanne Stadler, Alexander Meschtscherjakov, and Manfred Tscheligi. 2014. Towards Autonomous Cars: The Effect of Autonomy Levels on Acceptance and User Experience. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '14)*. ACM, New York, NY, USA, Article 11, 8 pages. DOI: http://dx.doi.org/10.1145/2667317.2667330

33. Julian B. Rotter. 1967. A new scale for the measurement of interpersonal trust. *Journal of Personality* 35, 4 (1967), 651–665. DOI: http://dx.doi.org/10.1111/j.1467-6494.1967.tb01454.x

34. SAE International. 2014. Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. (16 1 2014). http://standards.sae.org/j3016_201401/

35. Brandon Schoettle and Michael Sivak. 2014. *A Survey of Public Opinion about Autonomous and Self-Driving Vehicles in the U.S., the U.K., and Australia*. Technical Report UMTRI-2014-21. The University of Michigan Transportation Research Institute, Ann Arbor, MI, USA. http://hdl.handle.net/2027.42/108384

36. Marcel Walch, Tobias Sieber, Philipp Hock, Martin Baumann, and Michael Weber. 2016. Towards Cooperative Driving: Involving the Driver in an Autonomous Vehicle's Decision Making. In *Proceedings*

of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 16). ACM, New York, NY, USA, 261–268. DOI: http://dx.doi.org/10.1145/3003715.3005458

37. Adam Waytz, John Cacioppo, and Nicholas Epley. 2010. Who Sees Human?: The Stability and Importance of Individual Differences in Anthropomorphism. *Perspectives on Psychological Science* 5, 3 (2010), 219–232. DOI: http://dx.doi.org/10.1177/1745691610369336
38. Adam Waytz, Joy Heafner, and Nicholas Epley. 2014. The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology* 52 (May 2014), 113–117. DOI: http://dx.doi.org/10.1016/j.jesp.2014.01.005
39. Philipp Wintersberger, Andreas Riener, and Anna-Katharina Frison. 2016. Automated Driving System, Male, or Female Driver: Who'D You Prefer? Comparative Analysis of Passengers' Mental Conditions, Emotional States & Qualitative Feedback. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 16)*. ACM, New York, NY, USA, 51–58. DOI: http://dx.doi.org/10.1145/3003715.3005410
40. Nidzamuddin Md. Yusof, Juffrizal Karjanto, Jacques Terken, Frank Delbressine, Muhammad Zahir Hassan, and Matthias Rauterberg. 2016. The Exploration of Autonomous Vehicle Driving Styles: Preferred Longitudinal, Lateral, and Vertical Accelerations. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 16)*. ACM, New York, NY, USA, 245–252. DOI: http://dx.doi.org/10.1145/3003715.3005455