

ExplAI Yourself! Transparency for Positive UX in Autonomous Driving

Tobias Schneider
Stuttgart Media University
Stuttgart, Germany
schneider@hdm-stuttgart.de

Joana Hois
Mercedes-Benz AG
Stuttgart, Germany
joana.hois@daimler.com

Alischa Rosenstein
Mercedes-Benz AG
Stuttgart, Germany
alischa.rosenstein@daimler.com

Sabiha Ghellal
Stuttgart Media University
Stuttgart, Germany
ghellal@hdm-stuttgart.de

Dimitra Theofanou-Fülbier
Mercedes-Benz AG
Stuttgart, Germany
dimitra.theofanou-fuelbier@daimler.com

Ansgar Gerlicher
Stuttgart Media University
Stuttgart, Germany
gerlicher@hdm-stuttgart.de

ABSTRACT

In a fully autonomous driving situation, passengers hand over the steering control to a highly automated system. Autonomous driving behaviour may lead to confusion and negative user experience. When establishing such new technology, the user's acceptance and understanding are crucial factors regarding success and failure. Using a driving simulator and a mobile application, we evaluated if system transparency during and after the interaction can increase the user experience and subjective feeling of safety and control. We contribute an initial guideline for autonomous driving experience design, bringing together the areas of user experience, explainable artificial intelligence and autonomous driving. The AVAM questionnaire, UEQ-S and interviews show that explanations during or after the ride help turn a negative user experience into a neutral one, which might be due to the increased feeling of control. However, we did not detect an effect for combining explanations during and after the ride.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; *Empirical studies in visualization*; **Empirical studies in HCI**.

KEYWORDS

user experience, explainable artificial intelligence, autonomous driving

ACM Reference Format:

Tobias Schneider, Joana Hois, Alischa Rosenstein, Sabiha Ghellal, Dimitra Theofanou-Fülbier, and Ansgar Gerlicher. 2021. ExplAI Yourself! Transparency for Positive UX in Autonomous Driving. In *CHI Conference on Human Factors in Computing Systems (CHI '21)*, May 8–13, 2021, Yokohama, Japan. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3411764.3446647>



This work is licensed under a Creative Commons Attribution International 4.0 License.

CHI '21, May 8–13, 2021, Yokohama, Japan
© 2021 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-8096-6/21/05.
<https://doi.org/10.1145/3411764.3446647>

1 INTRODUCTION

In the case of autonomous driving, replacing the human driver with an artificial intelligence (AI) may lead to users that are sceptical towards the system. Subjective well-being is decreased and therefore, user experience (UX) is less positive. As Hartwich et al. demonstrated, passengers prefer human drivers over autonomous vehicles, even though the ride was identical [19]. The negative experiences may be, for example, due to users not being in control of the system, not being able to actively engage in the situation, disapproving the driving style, or not yet being familiar with autonomous vehicles (AVs) and their services [28, 38]. However, easy accessibility and practicability also show that users have an interest and curiosity in such new technology [44].

An in-depth survey is provided by Miller [39] on requirements for system explanations for humans. So far, only a few approaches have investigated whether available methods for explanations from AI systems can be directly provided to end-users [4, 5, 35]. Yet, the question is how to implement insights from human needs for explanation and how to comply with regulations for transparency for a specific use case that also benefits the system design with regard to UX. This is also of interest, as design guidelines for AI systems are available that recommend or even require systems to be transparent and explain themselves to developers, and end-users [26].

This paper focusses on vehicles with full driving automation (SAE level 5 [43]) and investigates if system explanations can improve the UX during autonomous rides. We conducted a user study to determine whether explanations of the system during or after an autonomous ride can improve the experience with the AV, or even support passengers' well-being. The focus was on first-time users or users who had only a few experiences with AVs so far. The study intersects with the following two fields of research: (1) User Experience Design (UXD) and interaction design which aim at providing a positive human-computer interaction and (2) Explainable Artificial Intelligence (XAI) that aims at increasing transparency in system behaviour and insights in system predictions.

2 RELATED WORK

This paper is related to the research topics UX, XAI, user trust and acceptance. It is applied in the context of first-time users or users

with little experience in autonomous driving. This paper is based on the following theoretical background and related to the following similar applications.

2.1 User Experience Design

When users encounter digital products, their interactions are complex, dynamic and subjective. UXD provides methods to create meaning and emotions during these product interactions [15, 23]. Designing for UX supports designers in creating product interactions and potentially associated emotions, even though emotions cannot be designed for directly due to their subjectiveness [8, 23]. This subjectiveness of UX can trigger feelings that cause users to continue or stop their current product interaction [32]. For the experimental setup below, the following models are used to evaluate the system interaction.

2.1.1 *The Model of Pragmatic and Hedonic Qualities.*

The model of pragmatic and hedonic qualities is a reductionist model of UX to measure the subjectiveness of UX by using a questionnaire [21]. It consolidates single experiences at different abstraction levels, so-called meta experiences, to analyse the pragmatic and hedonic dimension. The pragmatic dimension focuses on usability and ease of use to achieve intended goals during the interaction. The hedonic dimension focuses on the satisfaction of the users' well-being during the interaction based on the theory of universal psychological needs [21, 22]. The User Experience Questionnaire - Short (UEQ-S) [46] is a semantic differential questionnaire based on the model of pragmatic and hedonic qualities that provides a measurement for UX. It applies a seven-point Likert scale for eight items, four of which to measure the pragmatic and hedonic quality each.

2.1.2 *The Autonomous Vehicle Acceptance Model.*

Hewitt et al. [25] have proposed a model to assess the acceptance of AVs, the Autonomous Vehicle Acceptance Model Questionnaire (AVAM). The model aims at addressing user engagement and control when interacting with the AV, rather than technological aspects of vehicles and their degree of autonomy. The AVAM consists of 26 items and 3 interaction control aspects, measured on a seven-point Likert scale. Out of the 26 items, the following factors are measured: Performance Expectancy, Effort Expectancy, Social Influence, Facilitating Conditions, Attitude Towards (Using) Technology, Self-Efficacy, Anxiety, Behavioural Intention (to use the vehicle), and Perceived Safety.

2.2 Explainable AI

Explainable AI (XAI) [1, 10, 40] is concerned with methods to explain why an AI algorithm predicts a certain result. Explanations can be used for knowledge discovery, model verification, model improvement, but also user acceptance. For AI developers, XAI helps to evaluate and improve AI models. For end-users, it helps to clarify and improve the interaction with an AI system. With regard to machine learning methods, XAI techniques have been developed rather to analyse technical aspects than to present the explanations to end-users [11]. As Samek and Müller point out "the optimisation of explanations for optimal human usage is still a challenge which needs further study" [45, p. 17].

Technical explanation possibilities of an AI system may not necessarily be adequate or applicable for what end-users need in a specific situation [50]. Individual users have particular needs and preferences with regard to explanations under different conditions, which a system should take into account [39]. The combination of XAI and human reasoning is expected to benefit the human-AI-interaction [42] and individual needs during human reasoning can be addressed by different XAI methods [48]. XAI is one aspect to support transparency in the human-AI interaction [27]. Furthermore, transparency is also strongly recommended by different guidelines for AI system design [26, 30]. The European Commission's guidelines [26], for instance, emphasise that explanations from an AI systems should address the stakeholder concerns.

Connecting UXD and XAI to increase transparency in the human-AI interaction motivates this paper. Although no specific XAI methods or implementations are used directly, different explanation information types are evaluated for their effect on end-user interactions and UX. The results contribute to a better understanding when and how explanations can support the user interaction and potentially increase the UX.

2.3 Autonomous Vehicle Interaction

As of yet, the technical requirements to reach the level of autonomous or automated driving is an ongoing research topic [36]. Several aspects have to be taken into account for the interaction between human users and semi-automated vehicles [16]. Design recommendations are already available for semi-automated driving, e.g., handover scenarios [41]. For trust and user acceptance, transparency in the decision-making of the autonomous shuttles is essential [29].

Communicating AI decisions play an important role in autonomous driving. The study of Jeon et al. has shown that passengers want to be in control of the AV at any time [31]. This is particularly important in urban areas with a high number of simultaneous traffic situations, as this can lead to trust issues and the fear of automation failures [17]. An AV should thus communicate its awareness and intent to increase the feeling of control and security for passengers. UX and the feeling of trust in the system also increase when presenting additional information to passengers, such as head-up visualisations [24]. Trust is also increased in particularly difficult driving scenarios [49].

Besides internal communication between vehicles and passengers, external communication is also relevant for the communication with other traffic participants [9] and can have a positive effect on UX and pedestrians' attitude towards AVs [13].

Explaining the behaviour of autonomous systems can increase the user's understanding of their actions, and the overall transparency [27, 51]. Explanations of AVs before their acting can further improve the user trust [20].

When designing awareness and intent, designers should focus on providing so-called why-information [34]. One way of designing such a system can be feedback communicated through light [14, 37]. Another way to communicate awareness and intent is a visualisation with object recognition [12]. User trust in AVs has also been shown for driving performance and measured comfort



Figure 1: Exterior and interior of the driving simulator prototype.

[47], at least in the context of semi-autonomous vehicles. Furthermore, for visual communication text, icons, and augmented reality should be used together [18]. While trust in autonomous vehicles also increases over time [44] and passengers will most likely use their driving time for other activities than following the traffic and vehicle behaviour [33], we explicitly focus on first-time users or users with little experiences to facilitate their acceptance and UX.

3 EXPERIMENTAL SETUP

3.1 Prototype

We conducted the study in a static (non-moving) driving simulator that was equipped with four car seats (two in the front, two in the back), one frontal display, two lateral displays on each side, and lateral LED strip lights on the sides of the display and above the lateral sides. The simulation is intended to represent the concept of an autonomous (shared) shuttle. Participants entered the simulator from the side and took a seat within the enclosed driving space, see Figure 1. We constructed the shuttle as its own room to simulate a van-sized shuttle experience.

A pre-recorded driving scenario was displayed to the participants on all screens. We recorded the driving scene with three GoPro Hero 5 cameras (front, left, right), and added AR-simulated visualisations manually with Adobe After Effects and motion tracking. We assigned testers to two groups: Group A watched the original version of the scene recording. Group B watched the scene that was enhanced by *simultaneous explanatory information (live explanation)*:

- (1) overlay visualisations on the screen simulating augmented reality
- (2) textual information on the screen regarding velocity, speed limits, destination and time to reach the destination
- (3) lighting on the LED strip lights in different colours and the option to blink

These enhancements for individual driving situations and their added information to the scene are described in Table 1. After the ride, we provided both groups with information about the ride via a mobile application, which allowed for comparing the effect of live explanations versus delayed explanations on the UX. For participants who had no live explanations, the mobile application

was the only source of information on the ride and therefore the only way to better understand the shuttle’s behaviour (see Figure 3). The option to have a mobile application as a source of additional information after the ride might be especially helpful for first-time users or users, who were confused by the shuttle’s behaviour and were not provided with live explanations.

The live explanations were only shown on the front display. The lateral displays showed slightly blurred scene recordings to avoid distraction and foster a focus on the front display. The shuttle prototype did not provide any audio or car sounds, as the original audio file of the video recording had low quality due to wind noises.

3.2 Participant Groups

Overall, 40 participants were individually taking part in the experiment, 20 participants in each of the following groups:

- Group A : During the drive, participants get no live explanations. After the drive, participants are asked to interact with the drive summary on a mobile app.
- Group B : During the drive, participants get live explanations. After the drive, participants are asked to interact with the drive summary on a mobile app.

The gender distribution of participants was 35% female and 65% male (group A: 10 female, 10 male; group B: 4 female, 16 male). Their average age was 24.65 (SD=4.97). The group-specific age distribution for group A was 24.6 (SD=4.51) and for group B 26.1 (SD=5.43). The majority of participants primarily travels by public transport with 55%, 37.5% travel by car, 5% by bicycle, and 2.5% by foot. Based on their prior experiences with autonomous systems, participants are distributed in: 25% had no prior experiences, 50% had experiences with driving assistance such as cruise control, 12.5% had experiences with semi-automatic driving systems such as lane assists, 5% had experiences with highly automated driving systems such as highway and take-over driving assistants, and 7.5% had experiences with fully autonomous vehicles in research contexts. The participant group consisted of employees of a German automobile manufacturer, students and employees of a university as well as employees of a research campus. Each participant received a 10 Euro Amazon voucher for their participation.

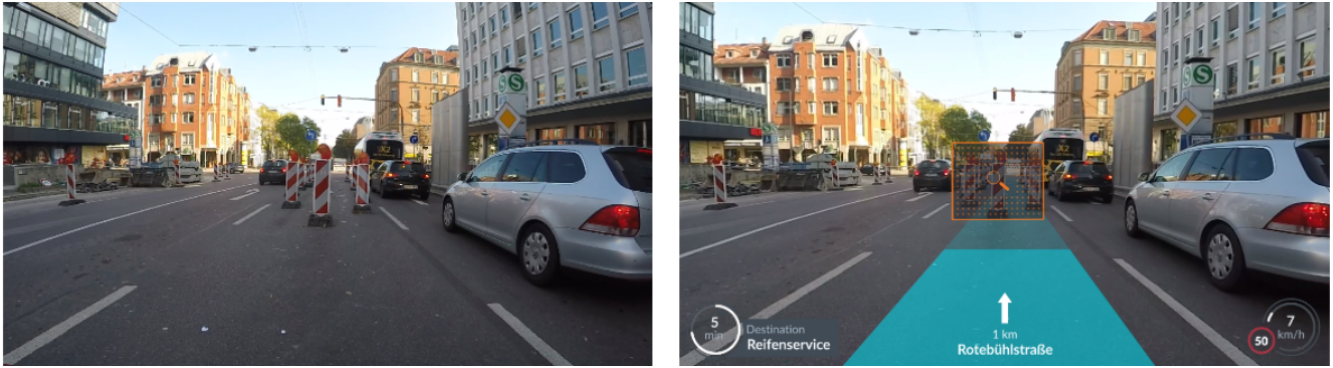


Figure 2: Driving scenario no. 1 as displayed to participants of group A (left) and group B (right).

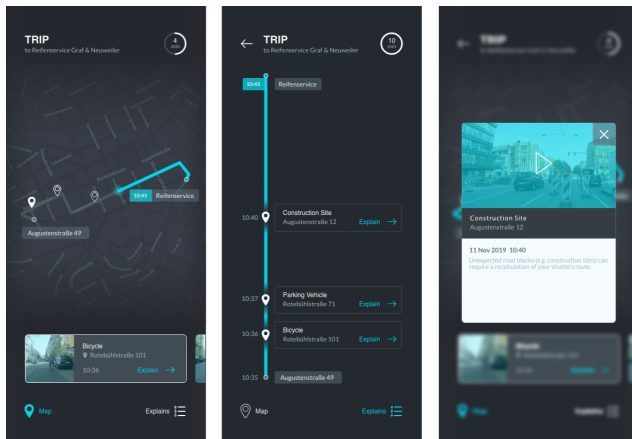


Figure 3: Different screens of the mobile application that was shown to the testers after the ride.

3.3 Explanations of Driving Situations in Shuttle Simulator and Mobile Application

For each participant, the experiment lasted 30 minutes on average. As part of the experiment, the overall driving time in the shuttle prototype was 8 minutes 25 seconds which is equivalent to an average taxi ride in Germany [3]. The recorded driving situation took place in an urban downtown area during the daytime. The maximum speed during the recording was 50 km/h, which was also the inner-city speed limit. The average velocity during the ride was 15 km/h, caused by the traffic flow, traffic lights, and obstacles.

During this driving time, eight driving situations were displayed that happened unexpectedly and caused the simulated vehicle to slow down and update its driving behaviour with regard to the traffic situation. The eight situations are listed in Table 1 together with the live explanations for group B.

We introduced participants to the experiment with the information that they are going to fill in questionnaires, take part in an autonomous driving experiment, and interact with a mobile app. Participants did not receive any additional tasks, e.g., about reaching their destination in a given time, to reduce potential stress

and to let participants focus on the ride itself. As participants regularly take part in public traffic situations, they were naturally able to judge the reaction of the vehicle during their ride, e.g., they were able to evaluate the vehicle's performance. During the ride, participants did not perform any other activities.

An example of driving scenario number 1 is shown in Figure 2. Only the participants of group B received live explanations in multiple ways. The LED lights were turned on and changed in colour. The front display highlighted the intended route by a blue transparent marking on the street. Detected objects in the street and their status were also marked on the display if they were relevant for the driving situation. The shuttle destination, the current road name, and the shuttle velocity were displayed on the bottom of the front display.

3.4 Experiment Design

Figure 4 shows the overall experimental sequence: (1) We introduced participants to the experiment and asked questions on demographic aspects, prior knowledge about AVs and their typical mobility behaviour. (2) Participants then entered the shuttle simulator and experienced the autonomous ride. Only participants of group B experienced the driving situations (Table 1) together with the live explanations. (3) After participants exited the shuttle simulator, they filled in the AVAM, UEQ-S, and one additional item regarding their subjective feeling of control, i.e., participants were asked to rate the statement “During the ride, I had the feeling to stay in control.” on a 7-point Likert-scale. This item was added, as the AVAM only refers to the user's control over body parts, which is less relevant for our fully autonomous context. According to Ajzen and Icek [2], the subjective feeling of control can be understood as a person's belief to be capable of obtaining a desired outcome and achieving a given goal. Therefore, a user should be able to experience a feeling of control even though they do not have actual control over the vehicles's behaviour during the ride or after using the mobile application. After the questionnaires, we also asked participants if they were feeling well or if anything was unclear and what else they would like to express about the ride. Since we focused on first-time users, there was no pre-drive evaluation of the participants UX. Instead, we emphasised a comparison between the evaluation directly after the ride and after the interaction with

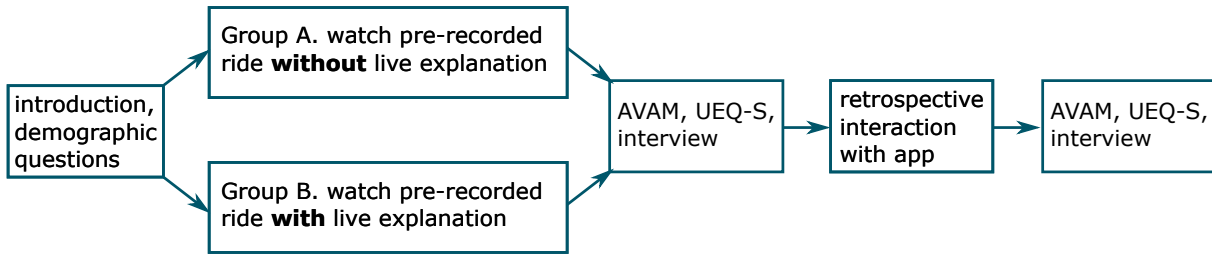


Figure 4: Experimental sequence for both participant groups.

Table 1: Driving situations and simultaneous explanatory information for participant group B.

No. (Minute)	Situation	Effect	Live Explanations shown to group B
1 (1:33-1:51)	a construction site is located in front of a traffic light and the lane that goes straight ends unexpectedly (cf. Figure 2)	the vehicle has to stop and update the route, as the road straight ahead is blocked and the lanes only allow to turn left or right	(1) the LED strip lights turn yellow, (2) the (simulated) AR visualisation highlights the obstacle in a yellow box, adjusts the size and reduces its colour intensity of the expected lane, visualises a processing status, displays ‘road changed’ with an ETA update, and updates the lane visualisation excluding the obstacle, (3) the vehicle continues to turn right
2 (3:37-3:48)	a pedestrian is crossing the street in front of the shuttle	the (already waiting) vehicle has to wait before it continues its drive	(1) the LED strip light that is closest to the pedestrian turns yellow, (2) the AR visualisation highlights the pedestrian with a yellow bounding box and adjusts the size and reduces its colour intensity of the expected lane, (3) after the pedestrian has left the road the bounding box disappears from the display
3 (3:51-4:56)	a bicycle drives rather slowly in front of the shuttle on the road	the vehicle has to slow down and is not able to safely pass the cyclist	(1) the initial red LED strip lights turn yellow and the cyclist is highlighted with a yellow bounding box on the front display, (2) the AR visualisation adjusts the size and reduces its colour intensity of the expected lane, (3) after the cyclist has left the road, the visual information disappears
4 (4:07-4:13)	a delivery person enters the road from behind a van and crosses the street	the vehicle has to stop abruptly	(1) the LED strip lights turn red, (2) the AR visualisation highlights the delivery person with a red bounding box and adjusts the size and reduces its colour intensity of the expected lane, (3) after the delivery person has left the road, the visualisations disappear
5 (4:14-4:20)	another vehicle takes the right of way	the vehicle has to slow down almost to a halt as the oncoming vehicle takes the right of way	(1) the LED strip lights turn yellow, (2) the AR visualisation highlights the oncoming vehicle with a yellow bounding box and adjusts the size and reduces its colour intensity of the expected lane, (3) after the oncoming vehicle has left the road, the visualisations disappear
6 (5:17-5:30)	a delivery person opens the driver’s door of the delivery truck on the opposite lane	the vehicle has to slow down and halt as the open door makes the lane too narrow to pass	(1) the LED strip light closest to the delivery person turns yellow, (2) the AR visualisation highlights the delivery person with a yellow bounding box and adjusts the size and reduces its colour intensity of the expected lane, (3) after the delivery person has closed the door, the visualisations disappear
7 (5:35-6:03)	another vehicle attempts to parallel park	the vehicle has to reverse to make room for the other vehicle to park	(1) the LED strip lights turn yellow, (2) the AR visualisation highlights the other vehicle with a yellow bounding box and adjusts the size and reduces its colour intensity of the expected lane, it shows a processing circle and informs about the resulting ‘backwards’ manoeuvre, (3) after the other vehicle has parked, the visualisations disappear
8 (6:38-6:42)	a driver is opening the door of their car that is parked parallel to the road	the vehicle has to slow down slightly and passes the driver at a distance	(1) the LED strip light closest to the driver turns yellow, (2) the AR visualisation highlights the driver with a yellow bounding box and adjusts the size and reduces its colour intensity of the expected lane, (3) after passing the driver, the visualisations disappear

the mobile application. (4) We then asked participants to interact with the app. It shows the (simulated) driven route and provides additional information about each driving scenario. Participants were asked to look at their ride again, and the different situations from Table 1. The location together with visual and textual information, which was already shown to group B on the frontal screen during the ride, was presented in the app. (5) Finally, participants filled in the AVAM, UEQ-S, including the adjusted control item and they were able to express their thoughts on the mobile app.

We chose a mixed-method within-subjects design since we studied whether the UX differs for live feedback during the ride and an app usage after the ride (retrospective feedback and its implications). Therefore, we did not let participants experience the ride twice (with and without live explanations) since this would have created a bias. This would have also undermined the purpose of the mobile application and therefore, not allowed us to find out if it is a useful method. For these reasons, we chose an independent control group without live explanations but retrospective explanations.

We chose the situations above to reflect driving situations that require an update of the planned ride as a reaction to the environment. Furthermore, they took place in an urban area to take the possibility of trust issues and fear of automation failures into account as described in [17]. The situations typically caused the shuttle to slow down or drive slowly. These situations reflect that the autonomous drive always prioritises safety. We expect this behaviour to cause impatience or confusion among the passengers, which is expected to lead to a decrease in their UX.

3.5 Hypotheses

The assumed hypotheses for the experimental setup are:

- H_{1.1}** *Hypothesis_{1.1} – Increased Perceived Feeling of Safety.* The vehicle often drives slowly to demonstrate a careful and safety-oriented driving behaviour. This is expected to increase a feeling of safety during the autonomous ride for passengers, and thus an increase in user experience. In this case, significant differences in AVAM question items ‘24. I believe that using the vehicle would be dangerous’, ‘25. I would feel safe while using the vehicle’, ‘26. I would trust the vehicle’ are expected between group A and B during the ride, as group B receives live explanations about the vehicle behaviour, and between group A during the ride and during the app usage, as the app provides additional information about the vehicle behaviour.
- H_{1.2}** *Hypothesis_{1.2} – Increased Understanding and Ease of Use.* Providing passengers with live explanations or with information after the ride about the AV’s actions will increase their understanding and ease of use of the shuttle. Significant differences are expected in AVAM question items ‘4. I would find the vehicle easy to use’, ‘5. My interaction with the vehicle would be clear and understandable’, ‘6. It would be easy for me to use the vehicle’ between group A and B as well as within group A during and after the ride while interacting with the app.
- H_{1.3}** *Hypothesis_{1.3} – Increased Interest.* Live explanations about the AV’s actions will make the autonomous ride more interesting for passengers. Significant differences are expected in AVAM

question items ‘14. The vehicle would make driving more interesting’, ‘15. Using the vehicle would be fun’.

- H₂** *Hypothesis₂ – Increased User Experience.* Providing passengers with live explanations during the ride or after the ride via a mobile app about the AV’s actions will increase their UX of the ride. Significant differences are expected in UEQ-S results for pragmatic quality (PQ) and hedonic quality (HQ).
- H₃** *Hypothesis₃ – Increased Perceived Feeling of Control.* Providing passengers with explanations during or after the ride will increase their perceived feeling of control. Participants were presented with a 7-point Likert scaled question item ‘During the ride, I had the feeling to stay in control’ and expect significant differences.

4 RESULTS

4.1 Autonomous Vehicle Acceptance Model Questionnaire

To verify the hypotheses H_{1.1}, H_{1.2} and H_{1.3} we used the AVAM [25]. The results of the groups were not normally distributed. Therefore, a Wilcoxon Signed Ranks test was performed for the dependent samples and a Mann-Whitney test for the independent samples instead of an ANOVA. We calculated the Cronbach Alpha [6] for all variables to see if they have an adequate internal consistency. Results with $p < 0.05$ are reported as significant. Since the questionnaire has 26 items and was used four times during the tests the focus lies on the questions relevant for the hypotheses.

4.1.1 Hypothesis_{1.1} – Increased Perceived Feeling of Safety. We grouped question item 24, 25 and 26 under the variable *perceived safety* by the AVAM. This variable does not show any significant changes between the groups, see Table 2. Hence, there is no evidence that supports hypothesis H_{1.1}.

Table 2: AVAM items 24, 25 and 26 combined to one variable (1–disagree, 7–agree). Group A was not presented with live explanations during the ride.

Perceived Safety (Items 24, 25, 26)				
Cronbach α : 0.871				
	A_{drive}	A_{app}	B_{drive}	B_{app}
values	5.23	5.30	5.45	5.62
SD	1.30	1.53	1.14	1.23
p	0.1865		0.136	
	A_{drive}	B_{drive}	A_{app}	B_{app}
values	5.23	5.45	5.30	5.62
SD	1.30	1.14	1.53	1.23
p	0.3405		0.3615	

4.1.2 Hypothesis_{1.2} – Increased Understanding and Ease of Use. Question items 4, 5, and 6 are grouped under the variable *effort expectancy* by the AVAM. The results are shown in Table 3.

We expected a positive effect of additional explanatory information on the perceived ease of use and the understanding of the system. It was found that the mobile application could significantly increase the perceived ease of use and the understanding of the

system. This, however, was only true for group A where no live explanations were provided ($A_{\text{drive}}=5.15$, $A_{\text{app}}=5.93$, $p=0.0025$). In the case of group B, who had live explanations, the mobile application used after the drive had no significant effect. Furthermore, there was a significant difference between the groups during the drive ($A_{\text{drive}}=5.15$, $B_{\text{drive}}=5.82$, $p=0.0495$). Overall, group B, where live explanations were provided, had higher ratings in ease of use and understanding of the system than group A. Regarding the mobile application, the group comparison was not significant.

Table 3: AVAM items 4, 5, and 6 combined to one variable (1–disagree, 7–agree). Group A was not presented with live explanations during the ride.

Effort Expectancy (Items 4, 5, 6)				
Cronbach α : 0.8				
	A_{drive}	A_{app}	B_{drive}	B_{app}
values	5.15	5.93	5.82	6.05
SD	1.19	1.12	0.86	0.81
p	0.0025		0.1175	
	A_{drive}	B_{drive}	A_{app}	B_{app}
values	5.15	5.82	5.93	6.05
SD	1.19	0.86	1.12	0.81
p	0.0495		0.4835	

4.1.3 Hypothesis_{1,3} – Increased Interest. The AVAM groups the question items 13, 14, and 15 under the variable *attitude towards using technology*. The results are shown in Table 4. None of the questions relevant to hypothesis $H_{1,3}$ showed significant changes. Hence, there is no evidence that supports hypothesis $H_{1,3}$ from the experiment.

Table 4: AVAM items 13, 14, and 15 combined to one variable (1–disagree, 7–agree). Group A was not presented with live explanations during the ride.

Attitude Towards Using Technology (Items 13, 14, 15)				
Cronbach α : 0.769				
	A_{drive}	A_{app}	B_{drive}	B_{app}
values	4.17	4.33	4.55	4.60
SD	1.38	1.45	1.44	1.38
p	0.251		0.312	
	A_{drive}	B_{drive}	A_{app}	B_{app}
values	4.17	4.55	4.33	4.60
SD	1.38	1.44	1.45	1.38
p	0.204		0.2485	

4.2 User Experience Questionnaire

We analysed the UEQ-S to verify the hypothesis H_2 – *Increased User Experience*. Figure 5 shows the results for the four different user test runs. The questionnaire reliability was measured using Cronbach alpha [6]. The UEQ-S defines values > 0.8 as a positive evaluation and values < -0.8 as a negative evaluation. Looking at the UEQ-S results, there is a strong indication that the hypothesis H_2 is valid.

Table 5: Results for feeling of control (1–disagree, 7–agree). Group A was not presented with additional explanatory information during the ride.

Item: During the ride I had the feeling to stay in control.				
	A_{drive}	A_{app}	B_{drive}	B_{app}
values	2.70	3.40	3.70	4.20
SD	1.69	1.35	1.34	1.51
p	0.0205		0.0255	
	A_{drive}	B_{drive}	A_{app}	B_{app}
values	2.70	3.70	3.40	4.20
SD	1.69	1.34	1.35	1.51
p	0.0175		0.044	

Table 6: Correlation between the feeling of control and the UEQ-S.

UEQ-S item	Pearson Corr.	Sig. (1-tailed)
obstructive – supportive	0.255*	0.011
complicated – easy	0.071	0.265
inefficient – efficient	0.219*	0.026
clear – confusing	0.328**	0.001
boring – exciting	0.359**	0.001
not interesting – interesting	0.341**	0.001
conventional – inventive	0.335**	0.001
usual – leading edge	0.333**	0.001

Participants rated the drive without live explanations (A_{drive}) with a neutral PQ ($M=-0.59$, $\alpha=0.67$), a negative HQ ($M=-1.41$, $\alpha=0.88$) leading to a negative overall score ($M=-1.00$). They rated the mobile application after the drive without live explanations (A_{app}) with a neutral PQ ($M=0.06$, $\alpha=0.55$) and a neutral HQ ($M=-0.26$, $\alpha=0.86$) leading to a neutral overall score ($M=-0.08$). For the drive with live explanations (B_{drive}) participants evaluated the PQ ($M=0.18$, $\alpha=0.89$) and the HQ as neutral ($M=-0.51$, $\alpha=0.71$) which resulted in a neutral overall score ($M=-0.17$). The mobile application after the drive with explanations (B_{app}) was rated with a neutral PQ ($M=0.30$, $\alpha=0.90$), a neutral HQ ($M=-0.50$, $\alpha=0.89$) and a neutral overall score ($M=-0.10$) by the participants. A Wilcoxon Signed Rank test revealed significant changes between A_{drive} and A_{app} ($p<0.001$). A Mann-Whitney test showed significant changes between A_{drive} and B_{drive} ($p<0.001$).

The UEQ-S results show that providing information on a mobile app after the drive without live explanations significantly increased the passengers' UX from negative to neutral. Using the application after the drive with additional explanatory information produced almost the same results and did not show an increased UX. Providing passengers with additional explanatory information during the drive, compared to no additional explanatory information, significantly increases their UX from negative to neutral. The results show that live explanations or explanations after the ride can increase the passengers' UX. However, as the results are still far from being above 0.8, the hypothesis is partly accepted.

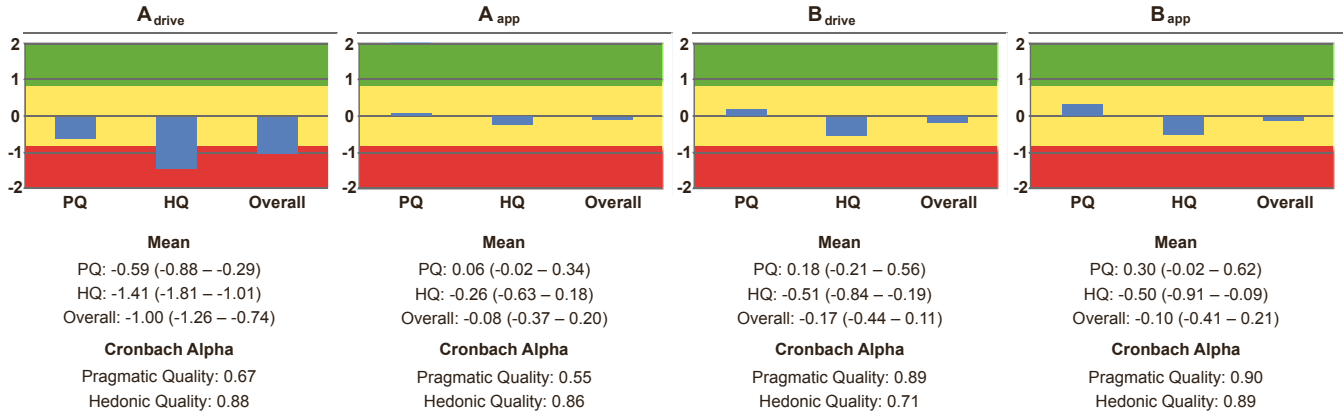


Figure 5: UEQ-S results of the different interactions during and after the ride for both participant groups. Values in brackets show lower to upper 95% confidence intervals.

4.3 Perceived Feeling of Control

All groups showed significant differences regarding the feeling of control, see Table 5. Providing an application after the ride significantly increases the feeling of control for both groups ($A_{drive}=2.70$, $A_{app}=3.40$, $p=0.0205$) ($B_{drive}=3.70$, $B_{app}=4.20$, $p=0.0255$), meaning that even passengers who received live explanations benefit from the mobile app. Providing passengers with live explanations during the ride also increases the feeling of control compared to the group that did not receive any information during the drive ($A_{drive}=2.70$, $B_{drive}=3.70$, $p=0.0175$). Hence, there is a strong indication that hypothesis H_3 – *Increased Feeling of Control* is valid.

Subsequently, we further analysed the relationship between the UX and the feeling of control (see Table 6). All word pairs, except *complicated* – *easy*, showed a weak to medium significant correlation to the feeling of control. As expected, we saw a positive correlation between a clear UX and feeling of control. However, the HQ dimensions are correlating stronger than the PQ. Therefore, a subjective feeling of control is a relevant factor when aiming for a positive UX.

4.4 Post Interview

As described in Table 4, we conducted interviews after the ride and after the app usage in addition to the questionnaires. We asked participants to describe how they felt during the ride, if there were situations in which something was unclear or in which they felt unwell. Participants were also asked how they would evaluate their system interaction particularly with regard to the shuttle's communication. They were also asked if they had anything else to add. Participants were able to freely articulate their thoughts during these interviews.

A summary of the transcribed and shortened statements from participants of both group A and B are shown in Table 7. The original statements were given in German. We clustered similar meanings together and referenced them by a shortened statement. The clustering was based on an inter-rater agreement of three raters on those terms. The overall numbers show how many of the participants have expressed the statement. It does certainly not indicate that other participants would have agreed or disagreed with the

statement. The statements should only be seen as complementary to the questionnaire results above.

The shuttle's driving is evaluated as slow from both groups. However, there is a distinction in the utterances if either the shuttle's driving itself is seen as slow or the time it took the shuttle to make a decision is seen as slow. The latter is mostly expressed by participants from group B, as they had access to the visualisations and the additional information from the shuttle during the ride. This indicates that they were more aware of the shuttle's actual manoeuvre planning, and it seems that the attention shifted from the shuttle's slow driving to the shuttle's slow decision making.

For 13 participants of group A, who did not get live explanations, the shuttle's driving reactions were not clear. 7 participants of group B shared this opinion. In contrast, the shuttle's driving reactions were clear to 12 participants of group B, and 11 participants of group B also stated that the live explanations on the screen were comprehensible. These statements support the increased HQ between A_{drive} and B_{drive} in the UEQ-S results as well as the significant changes in the AVAM effort expectancy shown in Table 3.

Only few participants commented that the ride was interesting (4 from group A, 2 from group B), 9 participants from group A and 3 from group B commented the ride to be boring. Similarly, we did not find a clear distinction for hypothesis $H_{1,3}$ with regard to AVAM's attitude towards using technology shown in Table 4.

In terms of perceived feeling of control, 5 people of group A stated that they did not feel in control compared to 3 of group B. These statements slightly support the significant result of Table 5 of hypothesis H_3 . A large number of participants stated to feel well during the ride, namely 10 of group A and 14 of group B.

When asked about the app usage, a larger number of participants from group A found the app to be helpful, whereas a larger number of participants of group B found the app to be unnecessary or unhelpful. This seems comprehensible as participants of group B already got the live explanations during the ride and had no interest in seeing them again after the ride. This adds to the differences between before and after app usage of group A and B, as described in section 4.2.

Table 7: Post interview answers of the participants grouped by statement.

Statement category.	Statements regarding the drive.	A _{drive}	B _{drive}
Shuttle manoeuvre	The situation detection is slow.	2	12
Shuttle manoeuvre	The shuttle drives cautiously and slowly.	8	3
Shuttle manoeuvre	A human would have reacted faster.	1	4
Shuttle manoeuvre	Interesting to see how the shuttle acts.	4	2
Shuttle manoeuvre	In some situations, the shuttle's reaction was unclear.	13	7
Shuttle manoeuvre	The shuttle's reactions were (always) clear.	5	12
Shuttle communication	The AR information is comprehensible.	–	11
Shuttle communication	The AR information was shown too late.	–	2
Participant feelings	I am not in control.	5	3
Participant feelings	The drive is boring.	9	3
Participant feelings	I was feeling well.	10	14
Statements regarding the app.		A _{app}	B _{app}
App	The app is unnecessary or not helpful.	6	13
App	The app is helpful.	14	6

5 DISCUSSION

Our goal was to investigate the effects of system explanations on the UX in an autonomous driving scenario. In order to do so, two experimental groups, one being provided with live explanations during the driving scenario and one having no explanations, were compared. Both groups, however, were able to interact with a mobile app that summarised the drive and explained the autonomous shuttle manoeuvres afterwards. As we expected, the UX is affected by the explanations that were provided during and after the drive. Group A, who had no live explanations, rated the UX significantly lower than group B with live explanations. Additionally, the mobile application afterwards was evaluated differently. While passengers with no live explanations seem to benefit from the mobile application after the drive, passengers who were not provided with live explanations do not. This finding, while preliminary, suggests that live explanations are sufficient to neutralise the possible negative effects autonomous driving has on the UX. The UEQ-S showed that the explanations helped to neutralise the negative effects on the HQ while it was possible to reach a positive PQ.

Our hypothesis regarding increased feeling of safety and the attitude towards using technology are not supported by the data. These results may be explained by the experimental setup of the study. As the experiment took place in a non-moving AV simulator, participants were continuously aware that they were not actually driven through real traffic by an autonomous system. A real driving scenario was only simulated. This may have influenced participants during their questionnaires and interview, particularly their estimation of feeling safe. This could explain the lack of significant results for hypothesis H_{1.1}. This explanation is supported by the fact that the rating for perceived safety was high among all groups and conditions and is backed by the qualitative results of the post-drive interviews. Regardless of the experimental group, a noteworthy number of people stated they have felt well during the drive. These comments might also reflect on insignificant results of the AVAM's perceived safety variable in Table 2. However, the large number of

statements is also a result of the interview, where participants were asked if they felt unwell or if anything was unclear.

Hypothesis H_{1.3} about increased interest in using the vehicle was also not supported by the data. The UEQ-S revealed that the dimension boring-interesting had the highest negative effect on the HQ among all groups and conditions. This might partly be explained by the fact that the driving scenario was only simulated and the driving was considered careful and slow by a large number of people. The results, however, implicate that creating an interesting experience for the user should be kept in focus when aiming for a positive UX. This might be achieved in a more tailored way of communicating with the passenger to ensure that only information of interest is provided. Therefore, a natural progression of this work is to analyse the effect of the user's preference as well as the impact of individual differences to allow for a more personalised communication. Additionally, ways of actively interacting with the AV or choosing and changing the communication and transparency settings might be helpful to avoid unnecessary explanations. This might be especially important in the long run, expecting the need for explanations to change when interacting with a system repeatedly. Since the focus of this study was on first-time users, more work will need to be done to determine the long-term effects of transparency on UX.

Regarding the feeling of control, the mobile application made a significant difference. This was true for both groups, i.e., the mobile application was beneficial even for passengers with live explanations. This is especially relevant since the subjective feeling of control was found to correlate positively with almost any dimension of the UEQ-S. This relationship should be evaluated further in future research.

The qualitative post-drive interviews largely support these results. Participants who had live explanations predominantly evaluated the communication as clear and comprehensive, whereas participants without live explanations stated more unclear situations and negative descriptions of the AV such as being boring and slow. The mobile application was described as unnecessary

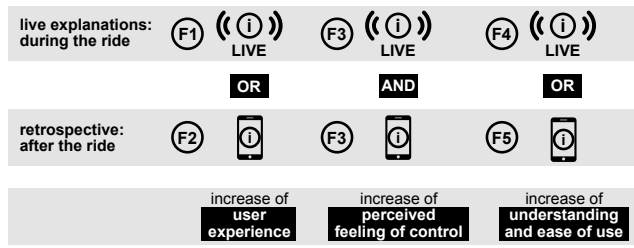


Figure 6: Annotated findings and contribution (numbers are broken down in section 7).

much more frequently by participants who were provided with live explanations. The evaluation of the app, however, might be influenced by the overall high feeling of safety during the ride. It is quite possible that more complex or more critical situations during the ride may increase the need to get in-depth explanations after the driving situation. This might be especially true in the case of a real driving scenario instead of a simulated one.

6 LIMITATIONS

As the experiment was conducted in a non-moving AV simulator, participants were continuously aware that they were not actually driving through real traffic by an autonomous vehicle. This may have influenced participants during their questionnaires and interview, particularly their estimation of feeling safe. This might explain the lack of significant results for hypothesis $H_{1.1}$. Recent studies estimate that more than 50% of asked consumers regard autonomous driving as being safe [7], which is in contrast to the rather high numbers (above 5 across all test groups) for perceived safety shown in Table 2. Another aspect that has to be taken into account are the rather young participants with 25 years on average, which might be a reason for their positive attitude towards AVs. Furthermore, the low age average does not represent a wider population, and thus, we did not analyse age-specific effects.

The simulator did also lack some aspects to support an immersive experience: no driving noise was available during the experiment, no visual depth was provided by the displays, and the display visualisations only simulated augmented reality. Also, as the visual enhancements on the displays were done manually, the intended route and marked items were perfectly highlighted without errors. In reality, visualisations would probably be less perfect.

With regard to the experimental task, participants were only “attending” the driving situation. They did not intend to travel from one location to another. If and how fast they reached the destination was of less importance to them, and they were also not instructed to pay attention to reaching the destination. As we focus on first-time users or users with little experience in autonomous driving, this is not a long-term study and participants were using the shuttle simulator only once during the experiment. When using the shuttle several times, however, the UX and interaction will most likely change while users become more familiar with the system.

Methodologically, the hypotheses could mostly be determined by the AVAM questionnaire variables. The additional hypothesis about a feeling of control, however, is not part of the AVAM and

participants were only asked to answer one item. In future experiments, we will extend items regarding control to evaluate if its significance prevails.

7 SUMMARY OF FINDINGS AND CONTRIBUTION

The main contribution of our paper is an initial guideline for autonomous driving experience design, bringing together the areas of UX, XAI and autonomous driving. These findings (F1-F5) may help other researchers to design for positive UX in autonomous driving (see annotated Figure 6).

- F1) Our study shows that possible negative effects of AD on the UX can be prevented by providing passengers of an AV with live explanations.
- F2) Retrospective explanations of the AV’s reaction to situations that happened during the ride, e.g. presented via a mobile app, show the same positive influence on the UX. However, they do not prevent any negative experiences during the ride. Furthermore, retrospective feedback does not show any additional positive effect on the UX if live explanations were provided.
- F3) Using live explanations and retrospective feedback, e.g. via a mobile application, positively influence the perceived feeling of control. Retrospective feedback shows a positive effect on the perceived feeling of control, even if live explanations were provided.
- F4) Providing live explanations during the ride increase the understanding and ease of use of the AV.
- F5) Retrospective explanations show the same effect if live explanations were not provided.

In addition, we found a correlation between the perceived feeling of control and the UX. However, at this point, we cannot find the exact reason. Do users have a better UX when they feel in control as they are able to understand the system better (even though when in fact they do not have any control)? In particular, more specific questionnaire items for this factor could lead to more insights. This will be further explored in a future mixed-method study.

8 CONCLUSION & OUTLOOK

In this paper we contribute an initial guideline for autonomous driving experience design, bringing together the areas of UX, XAI and autonomous driving. Our goal was to evaluate if a positive effect on UX can be achieved by explanations about the AV driving during and after the ride. Focussing on visual explanations seemed most adequate as the shuttle simulated the ride primarily on its screens. Using additional modalities, which are also related to driving, such as sound or seat vibration, could be tested further. Yet, it is to be expected that there is a threshold at which additional information confuses or overwhelms the passenger. In an early design phase, we had the idea to see if it makes a difference whether explanations are presented on the vehicle’s display or on a passenger’s personal mobile phone, i.e., if the use of a personal item may cause a priori trust. Thus, using personal, familiar display items versus non-familiar display items could be another experimental setup.

While the use of an AV simulator creates a better live experience than just a screen, it is still not a real driving situation. Putting

the test scenario into a driving simulator or a real traffic context could also be an option to pursue further. Our experiment focused solely on first-time users who did not interact with AVs before or users with little experience in that area. We expect that the need for explanations decreases over time when passengers are more used to the interaction. Assuming that passengers will use their driving time for other activities during their ride, the driving situation and its explanations may become less relevant. Longitudinal studies can lead to further insights.

Given that our research falls into the intersection of UX and XAI, there are more questions in general. One question is whether UX is more positive when using the exact algorithmic output from a visual recognition system on the screen or the more abstract (visual and verbal) descriptions on it. Providing the option to browse through different layers, from low-level feature analysis to high-level explanations, could be of interest to users and improve their technical understanding and potentially their UX and trust in the system, at least for certain user groups. Other user groups might prefer information that is further reduced, thus having a personalised way to let passengers choose their information channel is an interesting topic to pursue. Again, the effect of prolonged use, including critical situations, needs to be examined in further studies.

The acknowledgments section is defined using the "acks" environment (and NOT an unnumbered section). This ensures the proper identification of the section in the article metadata, and the consistent spelling of the heading.

ACKNOWLEDGMENTS

This work was carried out as part of the FlexCAR project and partially funded by the German Federal Ministry of Education and Research (funding number: 02P18Q647). It was executed in cooperation with the research campus Arena2036, the Stuttgart Media University and the Mercedes-Benz AG. The physical driving simulator build and the recording of the driving scenes were carried out by students of the Stuttgart Media University based on our specifications (in alphabetical order): P. Antony, F. Dimmler, F. Dums, L. Dutt, T. Hezel, A. Knizia, C. Knödler, B. Kramser, V. Lukina, M. Merz, N. Müller, K. Stängle, C. Zysk.

REFERENCES

- [1] Amina Adadi and Mohammed Berrada. 2018. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access* 6 (2018), 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
- [2] Icek Ajzen. 2002. Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior 1. *Journal of applied social psychology* 32, 4 (2002), 665–683.
- [3] Statistisches Amt. 2016. *Die wirtschaftliche Lage des Hamburger Taxengewerbes 2016*. Technical Report. <https://www.hamburg.de/taxi/2935760/taxigewerbe/>
- [4] Nadia Bekri, Jasmin Kling, and Marco Huber. 2019. A Study on Trust in Black Box Models and Post-hoc Explanations. 35–46. https://doi.org/10.1007/978-3-030-20055-8_4
- [5] Dawn Branley-Bell, Rebecca Whitworth, and Lynne Coventry. 2020. User Trust and Understanding of Explainable AI: Exploring Algorithm Visualisations and User Biases. In *Human-Computer Interaction. Human Values and Quality of Life*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 382–399. https://link.springer.com/content/pdf/10.1007%2F978-3-030-49065-2_27.pdf
- [6] Lee J. Cronbach. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 3 (Sept. 1951), 297–334.
- [7] Deloitte. 2019. *2019 Deloitte Global Automotive Consumer Study – Advanced vehicle technologies and multimodal transportation*. Retrieved June 08, 2020 from <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/manufacturing/us-global-automotive-consumer-study-2019.pdf>
- [8] John Dewey. 2005. *Art as Experience*. Penguin.
- [9] Debargha Dey, Andrii Matvienko, Melanie Berger, Bastian Pfleging, Marieke Martens, and Jacques Terken. 2020. Communicating the intention of an automated vehicle to pedestrians: The contributions of eHMI and vehicle behavior. *it - Information Technology* (2020). <https://doi.org/10.1515/itit-2020-0025>
- [10] Filip Došilović, Mario Brcic, and Nikica Hlupic. 2018. Explainable Artificial Intelligence: A Survey. In *2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE, 0210–0215. <https://doi.org/10.23919/MIPRO.2018.8400040>
- [11] Mengnan Du, Ninghao Liu, and Xia Hu. 2020. Techniques for Interpretable Machine Learning. *Commun. ACM* (2020), 68–77. <https://doi.org/10.1145/3359786>
- [12] Fredrick Ekman, Mikael Johansson, and Jana Sochor. 2016. To See or Not to See - The Effect of Object Recognition on Users' Trust in "Automated Vehicles". In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. ACM Press, New York, USA, 1–4.
- [13] Stefanie Faas, Lesley-Ann Mathis, and Martin Baumann. 2020. External HMI for self-driving vehicles: Which information shall be displayed? *Transportation Research Part F Traffic Psychology and Behaviour* 68 (01 2020), 171–186. <https://doi.org/10.1016/j.trf.2019.12.009>
- [14] Sarah Faltaous, Martin Baumann, Stefan Schneegass, and Lewis Chuang. 2018. Design Guidelines for Reliability Communication in Autonomous Vehicles. In *Proceedings of the 7th international conference on automotive user interfaces and interactive vehicular applications*. ACM Press, New York, New York, USA, 258–267.
- [15] Jodi Forlizzi and Katja Battarbee. 2004. Understanding Experience in Interactive Systems. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (Cambridge, MA, USA) (DIS '04). Association for Computing Machinery, New York, NY, USA, 261–268. <https://doi.org/10.1145/1013115.1013152>
- [16] Lex Fridman. 2018. Human-Centered Autonomous Vehicle Systems: Principles of Effective Shared Autonomy. *ArXiv* (2018). <https://arxiv.org/pdf/1810.01835.pdf>
- [17] Anna-Katharina Frison, Philipp Wintersberger, Tianjia Liu, and Andreas Riemer. 2019. Why do you like to drive automated? - a context-dependent analysis of highly automated driving to elaborate requirements for intelligent user interfaces. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. ACM Press, New York, USA, 528–537.
- [18] Peter Fröhlich, Raimund Schatz, Markus Buchta, Johann Schrammel, Stefan Suetter, and Manfred Tscheligi. 2019. "What's the Robo-Driver up to?" Requirements for Screen-based Awareness and Intent Communication in Autonomous Buses. *i-com* 18, 2 (2019), 151–165.
- [19] Franziska Hartwich, Cornelia Schmidt, Daniela Gräffing, and Josef F. Krems. 2020. In the passenger seat: differences in the perception of human vs. automated vehicle control and resulting HMI demands of users. In *International Conference on Human-Computer Interaction*. Springer, 31–45.
- [20] Jacob Haspiel, Na Du, Jill Meyerson, Lionel P. Robert Jr., Dawn Tilbury, X. Jessie Yang, and Anuj K. Pradhan. 2018. Explanations and Expectations: Trust Building in Automated Vehicles. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (Chicago, IL, USA) (HRI '18). Association for Computing Machinery, New York, NY, USA, 119–120. <https://doi.org/10.1145/3173386.3177057>
- [21] Marc Hassenzahl. 2007. The hedonic/pragmatic model of user experience. *Towards a UX manifesto* 10 (2007).
- [22] Marc Hassenzahl and Sarah Diefenbach. 2012. Well-being, need fulfillment, and Experience Design. In *Proceedings of the DIS 2012 Workshop on Designing Wellbeing*.
- [23] Marc Hassenzahl and Noam Tractinsky. 2006. User experience - a research agenda. *Behaviour & IT* 25, 2 (2006), 91–97.
- [24] Renate Häußlschmid, Max von Bülow, Bastian Pfleging, and Andreas Butz. 2017. Supporting Trust in Autonomous Driving. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces* (Limassol, Cyprus) (IUI '17). Association for Computing Machinery, New York, NY, USA, 319–329. <https://doi.org/10.1145/3025171.3025198>
- [25] Charlie Hewitt, Ioannis Politis, Theodoris Amanatidis, and Advait Sarkar. 2019. Assessing public perception of self-driving cars: The autonomous vehicle acceptance model. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. 518–527.
- [26] High-Level Expert Group on AI. 2019. *Ethics guidelines for trustworthy AI*. Technical Report. European Commission. Retrieved June 08, 2020 from <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- [27] Joana Hois, Dimitra Theofanou-Fuelbier, and Alischa J. Junk. 2019. How to Achieve Explainability and Transparency in Human AI Interaction. In *HCI International 2019 - Posters*. Springer, Cham, Cham, 177–183.
- [28] Lynn M. Hulse, Hui Xie, and Edwin R. Galea. 2018. Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety Science* 102 (2018), 1–13. <https://doi.org/10.1016/j.ssci.2017.10.001>
- [29] Calin Icolodean, Nicolae Cordos, and Bogdan Ovidiu Varga. 2020. Autonomous Shuttle Bus for Public Transportation: A Review. *Energies* 13, 11 (2020). <https://doi.org/10.3390/en13112917>

- [30] Information Commissioner's Office and The Alan Turing Institute. 2020. *Explaining decisions made with AI*. Retrieved June 08, 2020 from <https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence-1-0.pdf>
- [31] Myounghoon Jeon, Andreas Riemer, Jason Sterkenburg, Ju-Hwan Lee, Bruce N. Walker, and Ignacio Alvarez. 2018. An International Survey on Automated and Electric Vehicles: Austria, Germany, South Korea, and USA. In *Digital Human Modeling. Applications in Health, Safety, Ergonomics, and Risk Management*. Springer, Cham, Cham, 579–587.
- [32] Daniel Kahneman, Edward Diener, and Norbert Schwarz. 1999. *Well-being: Foundations of hedonic psychology*. Russell Sage Foundation.
- [33] Viktoriya Kolarova. 2020. Exploring the Elements and Determinants of the Value of Time for Manual Driving and Autonomous Driving using a Qualitative Approach. In *99th Annual Meeting of the Transport Research Board (TRB)*. <https://doi.org/10.1177/0361198120953770>
- [34] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. 2014. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJDeM)* 9, 4 (April 2014), 269–275.
- [35] Isaac Lage, Emily Chen, Jeffrey He, Menaka Narayanan, Been Kim, Samuel J. Gershman, and Finale Doshi-Velez. 2018. An Evaluation of the Human-Interpretability of Explanation. In *Conference on Neural Information Processing Systems (NeurIPS) Workshop on Correcting and Critiquing Trends in Machine Learning*.
- [36] Margarita Martínez-Díaz and Francesc Soriguera. 2018. Autonomous vehicles: theoretical and practical challenges. *Transportation Research Procedia* 33 (2018), 275–282. <https://doi.org/10.1016/j.trpro.2018.10.103> XIII Conference on Transport Engineering, CIT2018.
- [37] Bernt W. Meerbeek, Christel de Bakker, Yvonne A. W. de Kort, Evert J. van Loenen, and T. Bergman. 2016. Automated blinds with light feedback to increase occupant satisfaction and energy saving. *Building and Environment* 103 (July 2016), 70–85.
- [38] Johanna Meurer, Christina Pakusch, Gunnar Stevens, Dave Randall, and Volker Wulf. 2020. A Wizard of Oz Study on Passengers' Experiences of a Robo-Taxi Service in Real-Life Settings. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference (Eindhoven, Netherlands) (DIS '20)*. Association for Computing Machinery, New York, NY, USA, 1365–1377. <https://doi.org/10.1145/3357236.3395465>
- [39] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267 (Feb. 2019), 1–38.
- [40] Sina Mohseni, Niloofar Zarei, and Eric D. Ragan. 2018. A Survey of Evaluation Methods and Measures for Interpretable Machine Learning. *CoRR* abs/1811.11839 (2018). arXiv:1811.11839 <http://arxiv.org/abs/1811.11839>
- [41] National Highway Traffic Safety Administration. 2016. *Federal Automated Vehicles Policy: Accelerating the Next Revolution In Roadway Safety*. Retrieved June 24, 2020 from <https://www.hsdn.org/?abstract&id=795644>
- [42] P. Jonathon Phillips, Carina A. Hahn, Peter C. Fontana, David A. Broniatowski, and Mark A. Przybicki. 2020. *Four Principles of Explainable Artificial Intelligence*. Technical Report. <https://doi.org/10.6028/NIST.IR.8312-draft>
- [43] SAE On-Road Automated Vehicle Standards Committee and others. 2018. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. *SAE International Warrendale, PA, USA* (2018).
- [44] Arto O. Salonen and Noora Haavisto. 2019. Towards Autonomous Transportation. Passengers' Experiences, Perceptions and Feelings in a Driverless Shuttle Bus in Finland. *Sustainability* 11, 3 (2019). <https://doi.org/10.3390/su11030588>
- [45] Wojciech Samek and Klaus-Robert Müller. 2019. *Towards Explainable Artificial Intelligence*. Springer International Publishing, Cham, 5–22. https://doi.org/10.1007/978-3-030-28954-6_1
- [46] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence (IJIMAI)* 4, 6 (2017), 103–108.
- [47] Chris Schwarz, John Gaspar, and Timothy Brown. 2019. The effect of reliability on drivers' trust and behavior in conditional automation. 21 (2019), 41–54. Issue 21. <https://rd.springer.com/article/10.1007/s10111-018-0522-y>
- [48] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y. Lim. 2019. Designing Theory-Driven User-Centric Explainable AI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3290605.3300831>
- [49] Philipp Wintersberger, Tamara von Sawitzky, Anna-Katharina Frison, and Andreas Riemer. 2017. Traffic Augmentation as a Means to Increase Trust in Automated Driving Systems. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter (CHIItaly '17)*. Association for Computing Machinery, New York, USA. <https://doi.org/10.1145/3125571.3125600>
- [50] Christine T. Wolf. 2019. Explainability Scenarios: Towards Scenario-Based XAI Design. In *Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '19)*. Association for Computing Machinery, New York, USA, 252–257. <https://doi.org/10.1145/3301275.3302317>
- [51] Robert Wortham, Andreas Theodorou, and Joanna Bryson. 2016. What does the robot think? Transparency as a fundamental design requirement for intelligent systems. In *International Joint Conferences on Artificial Intelligence 2016: ethics for artificial intelligence workshop*.