Feedback Strategies for Crowded Intersections in Automated Traffic — A Desirable Future?

MARK COLLEY, Institute of Media Informatics, Ulm University, Germany
JULIAN BRITTEN, Institute of Media Informatics, Ulm University, Germany
SIMON DEMHARTER, Institute of Media Informatics, Ulm University, Germany
TOLGA HISIR, Institute of Media Informatics, Ulm University, Germany
ENRICO RUKZIO, Institute of Media Informatics, Ulm University, Germany

Automated vehicles should improve both traffic safety and user experience. While novel behavior patterns such as platooning become feasible to reduce fuel usage, such time- and fuel-reducing behavior at intersections can be perceived as unsafe and possibly disconcert users. Therefore, we designed and implemented nine feedback strategies for a simulated intersection and compared these in an online video-based between-subjects study (N=226). We found that visual feedback strategies limiting the view on the actual scene by providing calming views (a *Landscape* or the scene with hidden vehicles) were rated significantly higher in terms of perceived safety and trust. We discuss implications regarding future traffic and whether automated vehicles will necessitate altering reality for the user.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

ACM Reference Format:
Mark Colley, Julian Britten, Simon Demharter, Tolga Hisir, and Enrico Rukzio. 2022. Feedback Strategies for Crowded Intersections in Automated Traffic — A Desirable Future?. In *14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22), September 17–20, 2022, Seoul, Republic of Korea.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3543174.3545255

1 INTRODUCTION

Automated vehicles (AVs) are envisioned to reduce the number of accidents, fuel consumption, and travel time significantly [1]. These advancements are partially possible due to novel driving behavior such as platooning [19] or intersections that do not rely on traffic lights but can be handled through communication between the AVs [14]. However, despite these advantages, especially the automated intersections could be stressful for an AV user [28]. With the focus shift towards user experience in AVs [40], that is going away from driving-related to infotainment and working-related aspects [39], and the necessity to let the automation remain in control for these advantages to be possible, such stressful and potentially trust-diminishing events have to be considered for the introduction of AVs. Therefore, Krome et al. [28] performed a user study in which the participants encountered such an automated intersection from a first-person perspective. Results of this study showed that the trust in AVs within such traffic situations varies significantly [28]. This might be due to the lack of feedback from the AV to the user [28]. Therefore, we evaluated numerous feedback strategies for the automated intersection in a between-subjects online video-based study with *N*=226 participants. Our strategies are based on previous work targeted towards trust in AVs (e.g., [9, 31, 48]) and were activated approximately 100 m before the intersection. Participants encountered one concept from three different scenarios: on the middle lane

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2022 Copyright held by the owner/author(s).
Manuscript submitted to ACM

taking a left turn, on the outer lane driving straight across, and on the outer lane taking a left turn. Results indicated that the participants felt most comfortable when the reality was obfuscated. We discuss these results with respect to overtrust and reality loss aspects.

Contribution Statement: This work provides nine feedback strategies for the scenario of an automated intersection. In an online video-based between-subjects study with N=226 participants, we found that participants preferred to block out the potentially safety critical traffic instead of having full transparency. Our work opens a range of novel questions regarding the broader impact of AVs on future traffic and the role of feedback strategies.

2 RELATED WORK

This work builds on prior work in the domain of information visualization in AVs and trust calibration.

In the state of *calibrated trust* [32], the user's trust is appropriate to the automated system's capabilities. Calibrated trust is essential as it prevents issues with over- and undertrust. While Ekman et al. [16] emphasize that trust calibration happens before using an AV and is also formed after the journey, they also acknowledge the necessity for feedback during the ride. In this work, we focus on feedback communication.

Recent work poses the question of which information is necessary during an automated journey [9, 26, 30, 45, 49]. Wintersberger et al. interviewed 56 people, concluding "an inverse correlation between situational trust and participants' desire for feedback" [49, p.1]. They grouped their findings into six relevant themes for information communication: Predictability, impact, object characteristics, spatial properties, regulations, and visibility.

Wiegand et al. [45] identified 17 relevant situations for information presentation. Wiegand et al. conducted a think-aloud study based on these situations with 26 participants. The six main concerns were emotion and evaluation, interpretation and reason, vehicle capability, interaction, future driving prediction, and explanation request times [45]. Their scenarios mostly include other manual vehicles, misbehavior by the AV, or other vulnerable road users. They did not evaluate the use case of an automated intersection.

To support and calibrate trust, several approaches were investigated. These used various technical approaches and different levels of included uncertainty information. In their comparative work, Colley et al. [12] employed simulated Head-Up Displays (HUDs), Augmented Reality Windshield Display (WSD AR), and LED strips to investigate how AVs should communicate in critical situations. These critical situations included passing animals, children, and people with disabilities. These represent one-on-one (AV-on-person/animal) situations and are, therefore, fundamentally different from an automated intersection. The authors found that "no technologically advanced visualization (e.g., AR) is necessary for this visualization" [12, p. 17].

Other work used LED strips to inform the user of the AV's decisions via ambient light [31] or proposed to use light strips to show the intention or perception of other road users [46]. Lindemann et al. [29] employed an AR WSD to highlight potential threats (e.g., pedestrians or other vehicles). This resulted in higher situation awareness in low and high visibility scenarios than only displaying essential elements such as *speed* and *navigation info*.

In semi-automated vehicles explanatory information (i.e., *why* information) led to highest trust (compared to *how* information). Häuslschmid et al. [23] showed the vehicle's current situation interpretation via a world in miniature or via a simulated avatar of a chauffeur. The world in miniature increased trust most, however, no consensus among the participants could be found on whether such visualizations are necessary.

Currano et al. [13] tested an AR HUD and compared no HUD with a minimal and a complex one. Results were diverse and dependent on the dynamicity of the scene as well as the participants' reported driving style. Therefore, the authors concluded the necessity for an adaptive HUD. Regarding the future trajectory of the ego-vehicle, Schneider et al. [40]

evaluated explanations given via an AR WSD and a LED strip. They found that user experience for first-time users increased with the explanations which support the transferability of the results of Koo et al. [26]. The combination with an explanation after the journey via a smartphone app did not increase user experience.

The driving task is dependent on the encountered objects in the situation. For this, semantic segmentation is used [8]. Therefore, Colley et al. [10] investigated the effects of visualizations of the semantic segmentation task. This identification carries uncertainty. Thus, the authors argue that previous studies use abstract representations which do not enable the user to identify the cause for this uncertainty. Displaying the semantic segmentation inherently incorporates the uncertainty visualization. They found that their simulated AR WSD did not increase trust or mental workload. However, the subjective situation awareness was higher, and users rated recognition-related attributes significantly better. As the actions of an AV depend on other road users such as manual drivers or pedestrians, Colley et al. [9] compared pedestrian intention visualization in a VR study with 15 participants. They compared the visualizations on a tablet-based with an AR version simulating a WSD. The AR version reduced cognitive load significantly.

3 FEEDBACK STRATEGIES

This paper compares visualizations and auditory feedback. Therefore, we briefly introduce the design and rationale of each of the feedback strategies. All visualizations are dependent on an AR WSD as this was shown to be superior to tablet-based versions [9, 10] and represents the natural evolvement of HUDs. To better communicate that the shown visualizations on the AR WSD are not to be confused with the simulated reality, there is a blur at the edges of the AR WSD. The subtle blur was employed in all conditions and can best be seen in Figure 1e on the left side of the windshield.

The employed feedback strategies are based on previous work in the field of visualizations to support trust in AVs (e.g., [9, 10, 13, 40, 47]) and can be seen in Figure 1.

3.1 Arrows

We propose to use arrows (see Figure 1 b) to indicate the detection of other road users (also see [48]). We chose the color green to reflect that the vehicle has correctly detected them. The arrows move up and down slightly to better reflect the ongoing status of detection. We propose to visualize the detected objects in all views (i.e., windshield and peripheral or side windows).

3.2 Tinted Windshield

In this feedback strategy, the colour of the windshield and the windows on both sides of the vehicle are tinted to block the view to the outside (see Figure 1 c). For this to work in reality, smart switchable windows Oltean [33] could be used.

3.3 Landscape

With this feedback strategy, the existing environment is replaced by a forest (see Figure 1 d). The vehicle is still driving on a road that matches the real road. The idea was to create the illusion of the car driving through a calm environment to hide the intense situation while still keeping the road so the movement of the car in the "virtual" part of the scene matches the movement of the "real" part. The AV could use the road map data to generate this kind of virtual environment when necessary.

(a) *baseline*. (b) *Arrows*. (c) *Tinted Windows*.

(d) *Landscape*. (e) *Hide Vehicles*. (f) *Visualized Trajectories*.

(g) *LED stripe*. (h) *Outer Space*. (i) *Turning Seat*.

Fig. 1. The employed feedback strategies used in the experiment.

3.4 Hide Vehicles

For this feedback strategy, all other AVs around the user's car will be filtered out so that they are completely invisible while keeping the original environment (see Figure 1 e). Already, there is the possibility to remove vehicles from a video by using neural networks as shown by Zhang et al. [50].

3.5 Trajectory

Colley et al. [9] proposed to visualize pedestrian intention to calibrate user trust. Accordingly, the intention of the other AVs (i.e., their future trajectory) could be displayed for the user to indicate that the AV understands its environment. This is an extension to the *Arrows* concept as the displayed trajectories show **both** the recognition **and** the known intention. The trajectories are displayed in cyan because it has high conspicuity to other colors used in road traffic and is not associated with meanings such as red for warning or danger [44] (see Figure 1 f). We propose visualizing the trajectories in all views (i.e., windshield and peripheral or side windows).

3.6 LED Strips

Caberletti et al. [5] showed that lighting improved perceived safety significantly. While orange (compared to blue) was more comfortable and luxurious [5], orange is associated with danger and, thus, we opted for blue (see Figure 1 g).

3.7 Outer Space

Comparable to the *Tinted Windows*, *Hide Vehicles*, and *Landscape* visualizations, we designed a visualization altering the perceived reality by representing outer space (see Figure 1 h). This represents a futuristic and non-realistic display designed to provide novel user experiences. The designs *Tinted Windows*, *Hide Vehicles*, *Landscape*, and *Outer Space* can be viewed as visualizations on a spectrum for altering reality by simply blocking reality out by removing certain aspects (i.e., vehicles) and providing novel environments (from realistic: landscape to unrealistic: outer space).

3.8 Turning Seat

In future AVs, it is envisioned that the user can be seated in various directions and with several configurations [7]. While such new seating positions may raise some concerns in terms of safety, the study by Jin et al. [25] proposed a way of using a rotating seat to change the impact direction and improve crash safety. The Mercedes F015 [3] concept car already features turning front seats to allow for holding meetings during a car ride. A rotating driver's seat could also play a significant role in terms of safety. Similar to the PRE-SAFE® Impulse Side technology of the new Mercedes-Benz S-Class [4], in which the seat can move to the right in an accident from the left, the seat could turn away from the direction of the accident. In our concept, the user seat of the vehicle is rotated prior to the intersection, thereby, preventing the user from seeing the automated intersection (see Figure 1 i).

3.9 Voice Feedback

This concept builds upon the aspect of anthropomorphism by adding human-like features to the AV [38]. Especially the usage of voice was shown to significantly increase trust in AVs [20]. With this feedback strategy, an audio clip is played which contains a female voice produced with the text-to-speech feature by Google WaveNet E with the text "There is no need to be scared, the car has everything under control". This is done once before the intersection. The intention was to communicate to the user that the car has recognized the situation and is going to safely maneuver them through it.

4 EXPERIMENT

To evaluate the effects of different feedback strategies, we designed and conducted a between-subject study online video-based study with N=226. This study was guided by the research question (RQ): *What impact do the feedback strategies have on AV users in terms of (1) mental workload, (2) trust, (3) perceived safety, and (4) acceptance?*

4.1 Materials

To answer our research question, we designed three scenarios. We modeled the scenarios in Unity version 2020.2.1f1 [41]. The scene includes an intersection of two 3-lane roads in a large city. In the paper by Krome et al. [28], the scenarios included two-lane & four-lane roads with 400 & 800 vehicles crossing the intersection per lane per hour. For our implementation, we first opted for three lanes and 600 vehicles per lane per hour. However, as our goal was to induce the feeling of a crowded automated intersection with scenes resembling near misses, we decided to further increase the traffic density to 900 vehicles per lane per hour. 100 meters before arriving at the intersection, different types of feedback strategies are activated (see section 3). The intersection can be seen from the top in Figure 2. We used the model of a Tesla X and retracted the steering wheel to visually indicate that the vehicle is driving automatedly (see Figure 3). The overview of the course for each scenario is shown in Figure 4.

Fig. 2. The intersection was built in Unity with several vehicles crossing the intersection.

(a) Middle lane scenario. (b) Outer lane straight across scenario. (c) Outer lane driving left scenario.

Fig. 3. The three scenarios used in the experiment.

Fig. 4. The three different scenarios from above. 1 = Middle, 2 = Outer lane straight across , and 3 = Outer lane driving left scenario. The starting points of each route are located after the corresponding curve and the endpoints are set right after the intersection.

We used the same intersection but modeled three scenarios: on the middle lane taking a left turn, on the outer lane driving straight across, and on the outer lane taking a left turn. We opted to avoid including a right turn as taking a right turn does not involve risky behavior when only focusing on traffic without vulnerable road users such as bicyclists or pedestrians [24], which we focused on. The three scenarios can be seen in Figure 3. Including three scenarios allowed us to increase exposure time while varying the visual stimuli.

4.2 Measurements

We employed the mental workload subscale of the raw NASA-TLX [22] on a 20-point scale ("How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?"; 1=Very Low to 20=Very High). Additionally, we used the subscales *Predictability/Understandability* (*Understandability* from here) and *Trust* of the *Trust in Automation* questionnaire by Körber [27]. Understandability is measured using agreement on four statements ("The system state was always clear to me.", "I was able to understand why things happened."; two inverse: "The system reacts unpredictably.", "It's difficult to identify what the system will do next.") using 5-point Likert scales (1=*Strongly disagree* to 5=*Strongly agree*). Trust is measured via agreement on equal 5-point Likert scales on two statements ("I trust the system." and "I can rely on the system."). We also employed the van der Laan acceptance scale [42] with the subscales "usefulness" and "satisfying". Also, participants rated their perceived safety using four 7-point semantic differentials from -3 (anxious/agitated/unsafe/timid) to +3 (relaxed/calm/safe/confident) [17].
Finally, after experiencing all three scenarios, participants were able to provide open feedback.

4.3 Procedure

Every session started with an introduction, agreeing to the consent form, and a demographic questionnaire. Participants were instructed to use a laptop or a computer. We introduced participants as follows:
You are going to see videos of a highly automated ride through a city with highly dense traffic. You are approaching an intersection. All vehicles drive automatedly and communicate with each other in order to ensure that every car can safely cross the intersection. The AV is in control of the steering and accelerating/braking (lateral and longitudinal guidance). You are supposed to imagine sitting in such an AV, following the entire journey attentively, and then assessing it.

Participants were randomly assigned to one of the nine conditions or the baseline without their knowledge. Then they watched the videos of the feedback strategy they were assigned in all three scenarios in randomized order. After the video, participants answered the questionnaires detailed in subsection 4.2.

A script was running in the background that ensured window maximization and that participants could not skip or rerun the video (to ensure equal exposure time). A Full HD (1920 × 1080) monitor and loudspeakers were required. On average, a session lasted 12 min. Participants were compensated with 1.60€.

5 RESULTS

5.1 Data Analysis

Before every statistical test, we checked the required assumptions (normal distribution and homogeneity of variance assumption). In the case of non-normal data, a Kruskal-Wallis test was used for the concept comparison (between-subjects) and a Friedman's ANOVA was used to compare the scenarios (within-subject). The alpha level was 0.05.

For post-hoc tests, we used Bonferroni correction. For the figures, we used the R package ggstatsplot [35]. These include the mean or median (depending on the test), the density plots, the boxplots, the data points, and the post-hoc results. They also include statistical details (test used, number of observations, effect size, confidence interval). Therefore, we do not rewrite these in text. We used R 4.1.3 and RStudio 2022.02.0. All packages were up to date in April 2022.

5.2 Participants

We determined the required sample size via an a-priori power analysis using G*Power in version 3.1.9.7 [18]. To achieve a power of .8 with an alpha level of .05, 180 participants should result in an anticipated medium to high effect size (0.25 [21]) in a between-subjects design with ten groups and three measurements.

We recruited participants from the USA via https://www.prolific.co/. We restricted the participant pool was restricted to US citizens to avoid confounding effects of traffic handedness (right-hand vs. left-hand traffic) or culture [37]. Using an online participant database allowed us to circumvent biases found when recruiting mostly from a student population (e.g., see almost three-quarters of CHI publications in 2014 [6]).

N=226 participants (94 female, 129 male, and two non-binary) took part in the study. They were, on average, M=31.04 (SD=9.71) years old and 224 out of the 226 participants held a driving license on average of M=12.45 years. A driver's license was not required as AVs will, at least in certain operational domains, not need any intervention. A breakdown per condition is shown in Table 1.

Table 1. Demographics per condition. Kruskal-Wallis tests neither found a significant difference for age (p=0.15) nor for years of having a driver's license (p=0.71).

Condition	Number participants	Mean age	SD age	Mean license	SD license
Baseline	22	34.32	11.97	13.00	14.02
Arrows	25	30.92	10.12	12.12	10.44
Tinted	22	34.18	11.58	15.82	12.50
Landscape	22	30.73	11.16	11.41	11.40
Hide Vehicles	22	31.32	8.61	16.00	18.77
Trajectory	23	29.91	9.05	11.07	9.64
LED	22	27.73	7.60	10.14	8.07
Outer Space	25	31.28	7.20	11.24	8.57
Seat	23	27.48	7.43	9.65	7.71
Voice	20	34.50	10.33	15.20	11.22

70 of the 226 surveyed people use their car every day, 41 on working days, 60 participants drive three to four times a week, 14 participants once a week, 13 one to three times a month, and 27 out of 226 drive less than once in a month. 158 of the 226 participants have a university degree, 56 participants do not have a university degree but have a higher education entrance qualification, and nine have completed vocational training. 110 participants are employees, 33 are self-employed, 18 are unemployed, and 17 stated something else as a profession. We also had two pupils and 45 college students participating in our study. 70 respondents drive less than 7,000 km a year. 68 out of 226 drive between 7,000 and 14,999 km, 57 drive between 15,000 and 24,999 km, 24 between 25,000 and 32,999 km, and 14 even more than 33,000 km per year. On a 5-point Likert Scale (*1=Not at all, 5=Definitely*), participants reported a slightly high interest in AVs (M=4.15, SD=1.14). On average, the participants believe that automated driving can make their lives easier (M=4.10, SD=1.05) and that automated driving will become reality in the next 10 years until 2031 (M=4.26, SD=0.99).

5.3 Results of Feedback Strategies

5.3.1 Mental Workload and Trust. A Kruskal-Wallis test found no significant differences between the conditions for mental workload ($\chi^2(9)$=10.09, p=0.344). The lowest values were found for *Landscape* (M=7.70, SD=5.46) and the highest for *Tinted Windows* (M=9.95, SD=6.54).

A Kruskal-Wallis test found a significant difference between the conditions for Understandability (see Figure 5). Post-hoc tests found that subjective Understandability was significantly higher for the *Hide Vehicles* compared to *Tinted Windows* and *Outer Space*. Additionally, Understandability was significantly higher for *Landscape* compared to *Tinted Windows* and *Outer Space*.

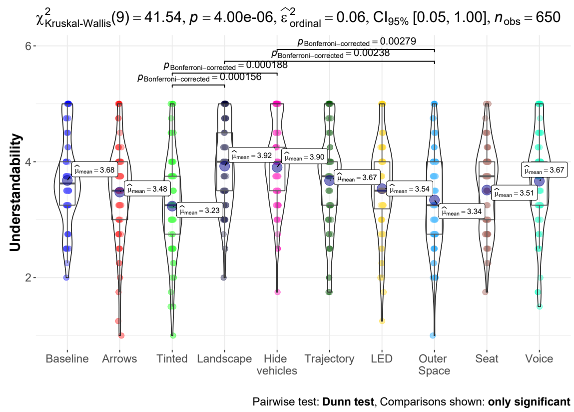

Fig. 5. Results of the Kruskal-Wallis test for Understandability.

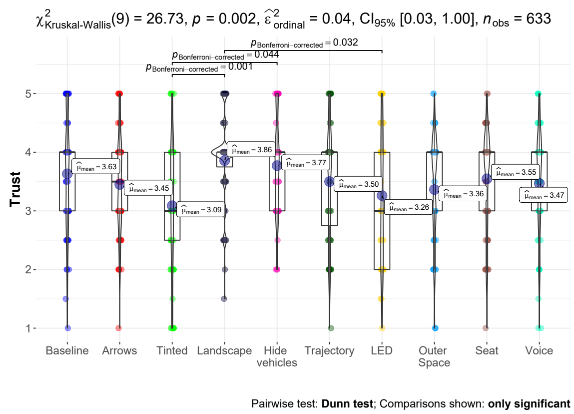

Fig. 6. Results of the Kruskal-Wallis test for trust.

A Kruskal-Wallis test also found a significant difference between the conditions for trust (see Figure 6). Post-hoc tests found that trust was significantly higher for *Landscape* than *Tinted Windows* and *LED stripe*, and that *Hide Vehicles* was rated significantly higher than *Tinted Windows*.

5.3.2 Usefulness, Satisfying, and Perceived Safety. A Kruskal-Wallis test found a significant difference between the conditions for usefulness (see Figure 7). Post-hoc tests found that usefulness was significantly higher in the *Landscape* condition compared to *Tinted Windows*, *Arrows*, and *Outer Space*. Also, *Hide Vehicles* was rated significantly higher than *Arrows*, *Tinted Windows*, *Outer Space*, and *LED stripe*. Finally, *Voice* was rated significantly higher than *Tinted Windows*.

A Kruskal-Wallis test found a significant difference between the conditions for satisfying (see Figure 8). Post-hoc tests found that the satisfying score was significantly higher in the *baseline* than the *Tinted Windows*. Additionally, the scores were significantly higher in the *Landscape* than *Arrows*, *Tinted Windows*, *LED stripe*, and *Outer Space*. Furthermore, *Hide Vehicles* was rated significantly higher than *Arrows*, *Tinted Windows*, *Visualized Trajectories*, *LED stripe*, *Outer Space*, and *Turning Seat*.

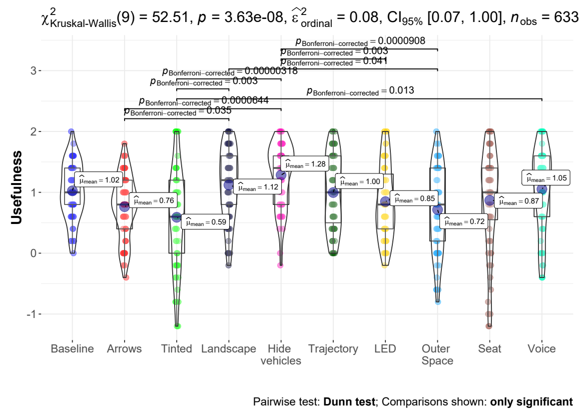

Fig. 7. Results of the Kruskal-Wallis test for usefulness.

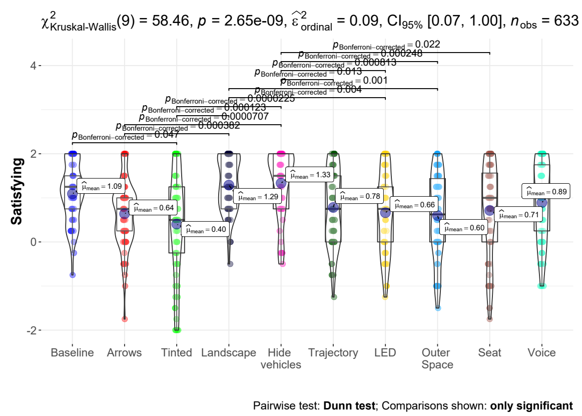

Fig. 8. Results of the Kruskal-Wallis test for satisfying.

A Kruskal-Wallis test found a significant difference between the conditions for perceived safety (see Figure 9). Post-hoc tests found that perceived safety was significantly higher in the *baseline* compared to the *Tinted Windows*. It was also significantly higher in the *Landscape* condition compared to *Arrows*, *Tinted Windows*, *LED stripe*, *Outer Space*, and *Voice*. Furthermore, *Hide Vehicles* was significantly safer than *Arrows*, *Tinted Windows*, *LED stripe*, *Outer Space*, and *Turning Seat*.

5.4 Results of Scenarios

For evaluating the scenarios, we eliminated data of eleven participants due to missing data points. Friedman's ANOVAs were employed as all scenarios were seen by every participant (thus, being within-subjects).

5.4.1 Mental Workload and Trust. A Friedman's ANOVA found no significant differences for mental workload ($\chi^2(2)$=2.94, p=0.230). Mental workload ranged from M=8.38 (SD=5.23; driving straight) to M=8.74 (SD=5.14; driving left in middle lane).

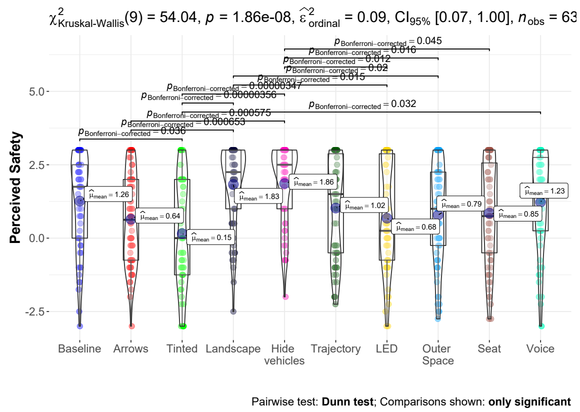

Fig. 9. Results of the Kruskal-Wallis test for Perceived Safety.

A Friedman's ANOVA found no significant differences for Understandability ($\chi^2(2)$=1.87, p=0.393). Understandability ranged from M=3.57 (SD=0.82; drive left outer lane) to M=3.63 (SD=0.76; drive left middle lane).

A Friedman's ANOVA found no significant differences for Trust ($\chi^2(2)$=0.81, p=0.667). Trust ranged from M=3.46 (SD=0.97; drive left outer lane) to M=3.53 (SD=0.97; drive straight).

5.4.2 Usefulness, Satisfying, and Perceived Safety. A Friedman's ANOVA found no significant differences for usefulness ($\chi^2(2)$=0.03, p=0.985). The mean value for usefulness was M=0.92 in all three videos.

A Friedman's ANOVA found no significant differences for satisfying ($\chi^2(2)$=5.60, p=0.061). Satisfying ranged from M=0.78 (SD=0.94; drive left outer lane) to M=0.87 (SD=0.93; drive straight).

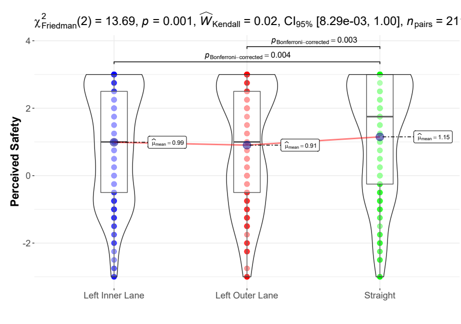

Fig. 10. Results of the Friedman's ANOVA for perceived safety for the videos.

A Friedman's ANOVA found significant differences for perceived safety. Post-hoc tests found that perceived safety was significantly higher in the video driving straight across than both videos driving to the left (see Figure 10).

5.5 Open Feedback

In general, the participants are interested in the topic but do not yet have complete confidence in AVs. One participant noted that the study videos were good but that the situation in the AVs continues to make him nervous. Another participant added that he could only imagine crossing an intersection if all cars were automated. If not, the participant would feel nervous as well. In addition, feedback was sent to us by mentioning errors that had previously occurred in AVs, which led to accidents. The participant noted that, despite everything, he was interested in the study but still liked being in control of the wheel. The only request made by a participant was for more transparency on what the AV is planning to do next. With our feedback strategy *Hide Vehicles*, for example, six out of seven feedback responses were positive. The participants added that there is a "big improvement in safe feeling", that the "self-driving car is a great idea" and that "it will prevent a lot of accidents". With the feedback strategy *Tinted Windows*, on the other hand, most of the feedback (11 out of 14) was negative. The complaint here was that "having the windows blacked out through the intersection made me feel very anxious. I think otherwise I would have enjoyed it much more." or that the participants "still worry about the reaction of these cars to unexpected events.".

6 DISCUSSION

Overall, all concepts and the *baseline* (except the *Tinted Windows* and *Outer Space*) had medium to high trust, usefulness, satisfying, and perceived safety ratings. Most of the significant differences could be determined between the *Hide Vehicles* and *Landscape* compared to the other concepts. However, we were not able to replicate findings regarding the positive effect of transparency [15, 26]. We discuss our findings in relation to the study design, the feedback concepts, and also with respect to the implications on traffic in general.

6.1 Relevance of Perspective

We found almost no differences and only one significant difference for the scenarios. Perceived safety was higher when driving straight through the intersection (see Figure 10). We attribute this to the fewer vehicles approaching the ego AV head-on. By repeated exposure to the intersection, we were able to reduce the bias of experiencing the automated intersection for the first time. Our results show that while most dependent variables were unaffected, perceived safety is likely easily altered by different behaviors, suggesting that studies on AV behavior should include multiple scenarios.

6.2 Necessity of Feedback in Automated Intersections

Krome et al. [28] stated that the in-car interface could drastically improve the users' experience. While we did not measure user experience directly, some feedback improved the perceived safety, usefulness, satisfying, and trust, the improvement was never significant compared to the *baseline*. It is unclear whether this is due to a too low participant number/too high number of conditions or more fundamental reasons. The data also does not align with previous work showing that transparency [15, 26] and AR systems [9, 23, 48] lead to increased trust, usefulness, satisfying, and perceived safety.

6.3 Transparency vs. Overlaying Reality

Previous work showed that transparent systems [15, 26] and AR systems [9, 23, 48] lead to increased trust, acceptance, and perceived safety. With our data, we were **not** able to show this effect. Especially the concepts *Arrows*, *Visualized Trajectories*, *LED stripe*, and *Voice* could be viewed as information presentation increasing transparency. However, these partially showed significantly lower ratings, for example, for trust (see Figure 6), usefulness (see Figure 7), and perceived safety (see Figure 9). The highest ratings were consistently reached for *Landscape* and *Hide Vehicles*, concepts that hide complexity rather than promote transparency. In the paper by Krome et al. [28], the authors claim the focus should be on finding ways to structure the overwhelming chaos of an automated intersection. Our data points toward not structuring but actually hiding the complexity. Regarding the technical feasibility, we argue that while this is not yet possible, the announcement of Panasonic [34], as well as the already available neural networks for removing other vehicles Zhang et al. [50], will allow such visualizations in the near future.

6.4 Driving in the Matrix: A Desirable Future?

Our results indicate that the participants trusted the AV more, felt safer, and had a more satisfying experience with the concepts *Landscape* and *Hide Vehicles*. This raises a fascinating and disturbing question: "Will automated vehicles and their novel possible behavior lead to the necessity to hide the reality to their users?" Currently, AVs are viewed from a lens of sheer endless possibilities: platooning [2] to reduce fuel usage, automated intersections as described by Krome et al. [28] and in this work, driverless taxis as already partially available by Waymo [43], the AV as the new living room [36], to name but a few. Mostly, the arguments for these novel behaviors are increased fuel efficiency and improved traffic flow [28]. As researchers and developers with a focus on the user experience in AVs, we must, therefore, look beyond these scenarios and critically reflect on which behavior is necessary and desirable. Also, if perfect throughput and fuel efficiency are desired, other options such as improved public transport should be focused on.

6.5 Limitations

By using simulated videos, we were able to address a wide participant group. However, as a consequence, external validity and transferability to the real world are limited. While we provided an immersive scenario and designed the scenarios realistically, the AV's vection could not be simulated with the chosen setup. Also, participants potentially did not perceive risks in critical situations due to the video setup. In future studies, this could be enhanced by using virtual reality (for a higher immersion) and simulators with higher degrees of freedom (e.g., [11]). Additionally, we solely focused on subjective dependent measures. Furthermore, our participant group was relatively young. Nonetheless, it was almost balanced based on gender. Finally, our concepts varied in some factors, for example, the employed modalities (only one using voice, the others using visual stimuli) or the required technical maturity (an *LED stripe* is easier to build than a *Tinted Windows* than the *Landscape* concept). While these designs are based on findings from the literature, an evaluation with distinct factors allowing a fully balanced study design could provide further insights.

7 CONCLUSION

In conclusion, we defined and compared nine feedback strategies designed to improve trust, mental workload, perceived usefulness, perceived satisfying, and perceived safety in an automated intersection. N=226 participants experienced one feedback strategy in the automated intersection from three scenarios: on the middle lane taking a left turn, on the outer lane driving straight across, and on the outer lane taking a left turn. Results indicate that the participants preferred to block out the potentially safety-critical traffic instead of having full transparency. Finally, we discuss these results in the broader light of future traffic.

ACKNOWLEDGMENTS

The authors thank all study participants. This work was conducted within the project 'SituWare' funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK).

REFERENCES

[1] Saeed Asadi Bagloee, Madjid Tavana, Mohsen Asadi, and Tracey Oliver. 2016. Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *Journal of modern transportation* 24, 4 (2016), 284–303.

[2] Soohyuk Bang and Soyoung Ahn. 2017. Platooning strategy for connected and autonomous vehicles: transition from light traffic. *Transportation Research Record* 2623, 1 (2017), 73–81.

[3] Mercedes Benz. 2015. The F015 Luxury in Motion. https://www.mercedes-benz.com/en/innovation/autonomous/research-vehicle-f-015-luxury-in-motion/. [Online; accessed 12-JUNE-2022].

[4] Mercedes Benz. 2020. 2021 Mercedes-Benz S-Class Safety Surpasses All Others. https://emercedesbenz.com/autos/mercedes-benz/s-class/2021-mercedes-benz-s-class-safety-surpasses-all-others/. [Online; accessed 12-JUNE-2022].

[5] Luca Caberletti, Kai Elfmann, M Kummel, and Christoph Schierz. 2010. Influence of ambient lighting in a vehicle interior on the driver's perceptions. *Lighting Research & Technology* 42, 3 (2010), 297–311.

[6] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 981–992. https://doi.org/10.1145/2858036.2858498

[7] Shi Cao, Pinyan Tang, and Xu Sun. 2020. Driver take-over reaction in autonomous vehicles with rotatable seats. *Safety* 6, 3 (2020), 34.

[8] Bowen Cheng, Maxwell D. Collins, Yukun Zhu, Ting Liu, Thomas S. Huang, Hartwig Adam, and Liang-Chieh Chen. 2020. Panoptic-DeepLab: A Simple, Strong, and Fast Baseline for Bottom-Up Panoptic Segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, New York, NY, USA, 12475–12485.

[9] Mark Colley, Christian Bräuner, Mirjam Lanzer, Walch Marcel, Martin Baumann, and Enrico Rukzio. 2020. Effect of Visualization of Pedestrian Intention Recognition on Trust and Cognitive Load. In *Proceedings of the 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '20)*. ACM, Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3409120.3410648

[10] Mark Colley, Benjamin Eder, Jan Ole Rixen, and Enrico Rukzio. 2021. Effects of Semantic Segmentation Visualization on Trust, Situation Awareness, and Cognitive Load in Highly Automated Vehicles. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 155, 11 pages. https://doi.org/10.1145/3411764.3445351

[11] Mark Colley, Pascal Jansen, Enrico Rukzio, and Jan Gugenheimer. 2022. SwiVR-Car-Seat: Exploring Vehicle Motion Effects on Interaction Quality in Virtual Reality Automated Driving Using a Motorized Swivel Seat. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 150 (dec 2022), 26 pages. https://doi.org/10.1145/3494968

[12] Mark Colley, Svenja Krauss, Mirjam Lanzer, and Enrico Rukzio. 2021. How Should Automated Vehicles Communicate Critical Situations? A Comparative Analysis of Visualization Concepts. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 94 (sep 2021), 23 pages. https://doi.org/10.1145/3478111

[13] Rebecca Currano, So Yeon Park, Dylan James Moore, Kent Lyons, and David Sirkin. 2021. Little Road Driving HUD: Heads-Up Display Complexity Influences Drivers' Perceptions of Automated Vehicles. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 511, 15 pages. https://doi.org/10.1145/3411764.3445575

[14] Kurt Dresner and Peter Stone. 2008. A multiagent approach to autonomous intersection management. *Journal of artificial intelligence research* 31 (2008), 591–656.

[15] Na Du, Jacob Haspiel, Qiaoning Zhang, Dawn Tilbury, Anuj K. Pradhan, X. Jessie Yang, and Lionel P. Robert. 2019. Look who's talking now: Implications of AV's explanations on driver's trust, AV preference, anxiety and mental workload. *Transportation Research Part C: Emerging Technologies* 104 (2019), 428–442. https://doi.org/10.1016/j.trc.2019.05.025 ID: 271729.

[16] Fredrick Ekman, Mikael Johansson, and Jana Sochor. 2017. Creating appropriate trust in automated vehicle systems: A framework for HMI design. *IEEE Transactions on Human-Machine Systems* 48, 1 (2017), 95–101.

[17] Stefanie M. Faas, Andrea C. Kao, and Martin Baumann. 2020. A Longitudinal Video Study on Communicating Status and Intent for Self-Driving Vehicle – Pedestrian Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376484

[18] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160.

[19] Pedro Fernandes and Urbano Nunes. 2012. Platooning with IVC-enabled autonomous vehicles: Strategies to mitigate communication delays, improve safety and traffic flow. *IEEE Transactions on Intelligent Transportation Systems* 13, 1 (2012), 91–106.

[20] Yannick Forster, Frederik Naujoks, and Alexandra Neukum. 2017. Increasing anthropomorphism and trust in automated driving functions by adding speech output. In *2017 IEEE intelligent vehicles symposium (IV)*. IEEE, Los Angeles, CA, USA, 365–372.

[21] David C. Funder and Daniel J. Ozer. 2019. Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science* 2, 2 (2019), 156–168. https://doi.org/10.1177/2515245919847202 arXiv:https://doi.org/10.1177/2515245919847202

[22] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, Amsterdam, The Netherlands, 139–183.

[23] Renate Häuslschmid, 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

[24] Kai Holländer, Mark Colley, Enrico Rukzio, and Andreas Butz. 2021. A Taxonomy of Vulnerable Road Users for HCI Based On A Systematic Literature Review. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 158, 13 pages. https://doi.org/10.1145/3411764.3445480

[25] Xin Jin, Haibin Hou, Ming Shen, Hequan Wu, and King H Yang. 2018. Occupant kinematics and biomechanics with rotatable seat in autonomous vehicle collision: a preliminary concept and strategy. In *Proceedings of the International Research Conference on the Biomechanics of Impact, Athens, Greece*. IRCOBI, Athens, Greece, 12–14.

[26] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. 2015. 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 (IJIDeM)* 9, 4 (2015), 269–275.

[27] Moritz Körber. 2019. Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*, Sebastiano Bagnara, Riccardo Tartaglia, Sara Albolino, Thomas Alexander, and Yushi Fujita (Eds.). Springer International Publishing, Cham, 13–30.

[28] Sven Krome, David Goedicke, Thomas J. Matarazzo, Zimeng Zhu, Zhenwei Zhang, J. D. Zamfirescu-Pereira, and Wendy Ju. 2019. How People Experience Autonomous Intersections: Taking a First-Person Perspective. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 275–283. https://doi.org/10.1145/3342197.3344520

[29] Patrick Lindemann, Tae-Young Lee, and Gerhard Rigoll. 2018. Catch my drift: Elevating situation awareness for highly automated driving with an explanatory windshield display user interface. *Multimodal Technologies and Interaction* 2, 4 (2018), 71.

[30] Andreas Löcken, Anna-Katharina Frison, Vanessa Fahn, Dominik Kreppold, Maximilian Götz, and Andreas Riener. 2020. Increasing User Experience and Trust in Automated Vehicles via an Ambient Light Display. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services* (Oldenburg, Germany) *(MobileHCI '20)*. Association for Computing Machinery, New York, NY, USA, Article 38, 10 pages. https://doi.org/10.1145/3379503.3403567

[31] Andreas Löcken, Wilko Heuten, and Susanne Boll. 2016. AutoAmbiCar: Using Ambient Light to Inform Drivers About Intentions of Their Automated Cars. In *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ann Arbor, MI, USA) *(AutomotiveUI '16 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 57–62. https://doi.org/10.1145/3004323.3004329

[32] Bonnie M Muir and Neville Moray. 1996. Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics* 39, 3 (1996), 429–460.

[33] Mihai Oltean. 2006. Switchable Glass: A Possible Medium for Evolvable Hardware. In *First NASA/ESA Conference on Adaptive Hardware and Systems (AHS'06)*, Vol. 2006. IEEE, Istanbul, Turkey, 81– 87. https://doi.org/10.1109/AHS.2006.69

[34] Panasonic. 2021. Panasonic Automotive Brings Expansive, Artificial Intelligence-Enhanced Situational Awareness to the Driver Experience with Augmented Reality Head-Up Display. https://na.panasonic.com/us/news/panasonic-automotive-brings-expansive-artificial-intelligence-enhanced-situational-awareness-driver. [Online; accessed 12-JUNE-2022].

[35] Indrajeet Patil. 2018. ggstatsplot: 'ggplot2' Based Plots with Statistical Details. https://doi.org/10.5281/zenodo.2074621

[36] Ingrid Pettersson, Annie Rydström, Helena Strömberg, Lena Hylving, Jonas Andersson, Maria Klingegård, and MariAnne Karlsson. 2016. Living Room on the Move: Autonomous Vehicles and Social Experiences. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (Gothenburg, Sweden) *(NordiCHI '16)*. Association for Computing Machinery, New York, NY, USA, Article 129, 3 pages. https://doi.org/10.1145/2971485.2987669

[37] Amir Rasouli and John K Tsotsos. 2019. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. *IEEE Transactions on Intelligent Transportation Systems* 21, 3 (2019), 900–918.

[38] Peter AM Ruijten, Jacques Terken, and Sanjeev N Chandramouli. 2018. Enhancing trust in autonomous vehicles through intelligent user interfaces that mimic human behavior. *Multimodal Technologies and Interaction* 2, 4 (2018), 62.

[39] Clemens Schartmüller and Andreas Riener. 2022. Multimodal Error Correction for Speech-to-Text in a Mobile Office Automated Vehicle: Results From a Remote Study. In *27th International Conference on Intelligent User Interfaces* (Helsinki, Finland) *(IUI '22)*. Association for Computing Machinery, New York, NY, USA, 496–505. https://doi.org/10.1145/3490099.3511131

[40] Tobias Schneider, Joana Hois, Alischa Rosenstein, Sabiha Ghellal, Dimitra Theofanou-Fülbier, and Ansgar R.S. Gerlicher. 2021. ExplAIn Yourself! Transparency for Positive UX in Autonomous Driving. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 161, 12 pages. https://doi.org/10.1145/3411764.3446647

[41] Unity Technologies. 2021. *Unity*. Unity Technologies.

[42] Jinke D Van Der Laan, Adriaan Heino, and Dick De Waard. 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies* 5, 1 (1997), 1–10.

[43] Waymo. 2020. Waymo is opening its fully driverless service to the general public in Phoenix. https://blog.waymo.com/2020/10/waymo-is-opening-its-fully-driverless.html. [Online; accessed 12-JUNE-2022].

[44] Annette Werner. 2018. New Colours for Autonomous Driving: An Evaluation of Chromaticities for the External Lighting Equipment of Autonomous Vehicles. https://doi.org/10.25538/tct.v0i1.692 Number: 1.

[45] Gesa Wiegand, Malin Eiband, Maximilian Haubelt, and Heinrich Hussmann. 2020. "I'd like an Explanation for That!"Exploring Reactions to Unexpected Autonomous Driving. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services* (Oldenburg, Germany) *(MobileHCI '20)*. Association for Computing Machinery, New York, NY, USA, Article 36, 11 pages. https://doi.org/10.1145/3379503.3403554

[46] Marc Wilbrink, Anna Schieben, and Michael Oehl. 2020. Reflecting the Automated Vehicle's Perception and Intention: Light-Based Interaction Approaches for on-Board HMI in Highly Automated Vehicles. In *Proceedings of the 25th International Conference on Intelligent User Interfaces Companion* (Cagliari, Italy) *(IUI '20)*. Association for Computing Machinery, New York, NY, USA, 105–107. https://doi.org/10.1145/3379336.3381502

[47] Scott R Winter, Stephen Rice, Nadine K Ragbir, Bradley S Baugh, Mattie N Milner, Bee-Ling Lim, John Capps, and E Anania. 2019. *Assessing pedestrians' perceptions and willingness to interact with autonomous vehicles*. Technical Report. US Department of Transportation. Center for Advanced Transportation Mobility ....

[48] Philipp Wintersberger, Anna-Katharina Frison, Andreas Riener, and Tamara von Sawitzky. 2019. Fostering User Acceptance and Trust in Fully Automated Vehicles: Evaluating the Potential of Augmented Reality. *PRESENCE: Virtual and Augmented Reality* 27, 1 (2019), 46–62. https://doi.org/10.1162/pres_a_00320 arXiv:https://doi.org/10.1162/pres_a_00320

[49] Philipp Wintersberger, Hannah Nicklas, Thomas Martlbauer, Stephan Hammer, and Andreas Riener. 2020. Explainable Automation: Personalized and Adaptive UIs to Foster Trust and Understanding of Driving Automation Systems. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) *(AutomotiveUI '20)*. Association for Computing Machinery, New York, NY, USA, 252–261. https://doi.org/10.1145/3409120.3410659

[50] Rong Zhang, Wei Li, Peng Wang, Chenye Guan, Jin Fang, Yuhang Song, Jinhui Yu, Baoquan Chen, Weiwei Xu, and Ruigang Yang. 2020. AutoRemover: Automatic object removal for autonomous driving videos. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. PKP Publishing Services, CA, USA, 12853–12861.