A Real Bottleneck Scenario with a Wizard of Oz Automated Vehicle - Role of eHMIs

Hatice Şahin Ippoliti
hatice.sahin128@gmail.com
Media Informatics and Multimedia Systems, University Oldenburg
Oldenburg, Germany

Angelique Daudrich
angelique.daudrich@uni-oldenburg.de
Media Informatics and Multimedia Systems, University Oldenburg
Oldenburg, Germany

Debargha Dey
debargha.dey@cornell.edu
Information Sciences, Cornell Tech
New York, United States

Philipp Wintersberger
philippwintersberger@gmail.com
Department of Digital Media, University of Applied Sciences Upper Austria
Hagenberg, Austria

Shadan Sadeghian
shadan.sadeghian@uni-siegen.de
University of Siegen
Siegen, Germany

Susanne Boll
susanne.boll@informatik.uni-oldenburg.de
Media Informatics and Multimedia Systems, University of Oldenburg
Oldenburg, Germany

Figure 1: Ghost driver wearing car seat costume and sitting behind the wheel of Wizard of Oz Automated Vehicle

AutomotiveUI '23, September 18–22, 2023, Ingolstadt, Germany

ACM ISBN 979-8-4007-0105-4/23/09…$15.00
https://doi.org/10.1145/3580585.3607173

ABSTRACT

Automated vehicles (AVs) are expected to encounter various ambiguous space-sharing conflicts in urban traffic. Bottleneck scenarios, where one of the parts needs to resolve the conflict by yielding priority to the other, could be utilized as a representative ambiguous scenario to understand human behavior in experimental settings. We conducted a controlled field experiment with a Wizard of Oz automated car in a bottleneck scenario. 24 participants attended the study by driving their own cars. They made yielding, or priority-taking decisions based on implicit and explicit locomotion cues on AV realized with an external display. Results indicate that acceleration and deceleration cues affected participants' driving choices and their perception regarding the social behavior of AV, which further serve as ecological validation of related simulation studies.

CCS CONCEPTS

• **Human-centered computing → Displays and imagers**; **Field studies**; **User studies**; **Interface design prototyping**.

KEYWORDS

Wizard of Oz, Automated Vehicle, AV, AV - Driver Interaction, eHMI, External Human- Machine Interfaces, field study, game of chicken, bottleneck, prosocial, trust, interview, mixed methods

ACM Reference Format:

Hatice Şahin Ippoliti, Angelique Daudrich, Debargha Dey, Philipp Wintersberger, Shadan Sadeghian, and Susanne Boll. 2023. A Real Bottleneck Scenario with a Wizard of Oz Automated Vehicle - Role of eHMIs. In *15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '23), September 18–22, 2023, Ingolstadt, Germany.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3580585.3607173

1 INTRODUCTION

Imagine yourself driving your regular car and approaching a narrow passage where two cars are parked on both sides of the road. On the opposite side, a passengerless self-driving vehicle is approaching the narrowing almost at the same time as you are. Who is going to take priority, and why? Even when fully automated vehicles will become available, there will be a longer phase of "mixed traffic", where both automated and manually driven vehicles are present. Game of chicken scenarios with automated vehicles (AVs) as in this example pose a crucial consideration for the usability of the services of AVs, such as, how their intention will be understood by humans, and if humans change their social behavior around them. In ambiguous scenarios, humans rely on informal communication originating from vehicle locomotion cues such as acceleration, and from humans, such as gestures and eye contact. While previous simulation studies suggest that vehicle locomotion cues may suffice to resolve the conflicts in bottleneck scenarios [46], an explicit indication of locomotion intention helps to resolve the conflicts more efficiently [57].

We validated these insights on a controlled test track study by using the Wizard of Oz (ghost driver) method [62]. In our study, we tested explicit locomotion intention cues conveyed with an external display in a realistic bottleneck scenario. Thereby, we hypothesized that in this ambiguous driving scenario, indicating locomotion intention cues with an external interface will alter (1) the driving choices of participants, (2) social perception of AV, and (3) trust in AV compared to non-indication of locomotion intention cues.

2 BACKGROUND

This section presents general AV-human interaction dynamics and conflicts, assistive role of external communication interfaces with an emphasis on bottleneck scenarios. It also reviews social interaction dynamics and the role of trust in AV-human interaction.

2.1 AV-Human interaction

Markkula et al. [41] define interactions as a space-sharing conflict in traffic or an "event with a collision course where interactive behavior is a precondition to avoid an accident". Some interactions can be resolved seamlessly by following predetermined traffic rules, while others, where the rules or the intentions of other road user(s) are unclear, require special communication between road users [18]. In such ambiguous situations, communication between both road users is particularly important. For example, Risto et al. [60] showed that pedestrians rely particularly on communication with drivers when crossing roads. If communication is missing or misunderstood, it can lead to the participants feeling uncomfortable in the crossing situation, or in the worst case, conflicts. This raises the question: how do road users decide who goes first and how is this communicated?

Communication between different traffic participants can be classified as explicit and implicit: explicit communication, such as simple gestures of waving, nodding, smiling, or blinking with car lights, can be helpful to signal the intention to other road users [16, 50, 70]. For example, in the study by Myers et al. [51], more drivers yielded when pedestrians showed hand signals compared to no signals. Moreover, early research in the field of AV-pedestrian interaction centered around the assumption of the critical importance of eye contact. Indeed, eye contact was shown to help establish communication and situation awareness between driver and pedestrian [26, 40, 54]. However, as the research matured, the critical importance of explicit communication and eye contact in normal, unambiguous situations was largely disproved. Numerous subsequent (more recent) studies highlight that explicit communication and eye contact are not as important a factor in crossing decisions, and that vehicle kinematics (movement patterns) play a much more important role in vehicle-pedestrian interaction [3, 16, 36, 49]. Šucha et al. [65] also corroborated the importance of vehicle kinematics by showing that the speed of the vehicle is a particularly important determinant of crossing willingness in pedestrians, along with the distance of the car, traffic density, and direction of approach. However, there is evidence that when the situation is ambiguous and the intention of the vehicle is not clear, pedestrians resort to explicit communication to seek confirmation [16]. In the case of road users interacting with Automated Vehicles (AVs) in such ambiguous situations where the intention of the vehicle is not clear from its kinematics alone, the lack of explicit, driver-centric communication poses a problem.

To solve the communication problem of missing driver-centric communication in AVs, eHMIs (External Human-Machine Interfaces) have been proposed. eHMIs are visual or auditory interventions geared towards communicating with other road users to disambiguate the intention of an AV. Various eHMI concepts have been proposed to date, and the design space for eHMIs is filled with different implementations which vary in terms of modality, placement, message, and many other dimensions [13]. Although there is no consensus with regard to the nature of eHMIs, numerous studies that have evaluated eHMIs in various forms have shown them to be effective solutions [1, 11, 12, 20, 27, 30]. Furthermore, in their real-world AV-pedestrian interaction experiment, Dey et al. [15] showed that when the intention of the vehicle is clear from its movement/kinematics alone, pedestrians do not need an eHMI, but it can disambiguate situations where the intention of the vehicle is unclear from vehicle kinematics.

2.2 Bottleneck/Game of Chicken Scenarios and eHMI

One of the classic scenarios of ambiguity, when AVs interact with Manually-Driven Vehicles (MDVs – non-automated, ordinary cars), arises when there is a deadlock or bottleneck situation without any clear rules that dictate the right of way, and there is a clear need for initiating a communication within a comfortable window of time for a seamless interaction [56]. Previous work has shown that vehicle kinematics still have a significant role to play here in communicating intent – lateral movement within the road (driving close to the edge of the road vs. occupying more road space by driving in the center) offers a very clear indication of intent as opposed to longitudinal movements (speed) [46, 58]. Furthermore, drivers of MDVs expect AVs to yield, and complying AVs were perceived as more trustworthy [45], although novel behaviors from AVs can confuse drivers [11]. That said, eHMIs are shown to increase perceived safety and reduce mental workload [11], while also facilitating shorter passing time and reducing crashes [55, 59]. A recent simulation study also corroborated these insights and found that an ideal way for AVs to communicate intent in bottleneck situations is by a combination of eHMI and employing lateral movements [57]. Prior research has also shown that eHMIs are not universally beneficial, and that they can have adverse effects in terms of overtrust [30] and violation of safety arising from confusion or ambiguity [22]. However, the substantial corpus of research showcasing the potential advantages of eHMIs outweigh the drawbacks, and we argue that this warrants a real-world investigation to evaluate its ecological validity, especially in this context of bottleneck negotiation.

From the social interaction perspective, these bottleneck or the game of chicken [53] scenarios in which one of the road users makes the decisive move to insist on the right of way or to "chicken out", humans may adapt their behavior in favor of themselves if AVs are strictly defensive and conflict-avoidant [7, 21, 44]. This supposition is supported by insights from recent studies: in a large survey conducted across China and South Korea, Liu et al. [38] found that individuals had an increased intention to bully AVs compared to human drivers, and they drew attention to potential hindrances of the deployment of AVs due to the aggressive or antagonistic behavior of humans, a phenomenon also corroborated in a study conducted in the United States [48]. While past research has investigated the function of eHMIs in terms of courteous behavior and polite strategies [34] and prosocial behavior by means of perceived traffic climate [64], the potential of eHMIs for improving cooperative and positive behavior in traffic remains unexplored.

While keeping these social dynamics in mind, previous research has established that trust in automation is the key factor for interacting with them and resolving conflicts. Trust in automation can be defined as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" [35]. Due to a series of trust-related accidents with Tesla Autopilot, the psychological construct of trust has become one of the key issues that need to be resolved to allow a successful implementation of AVs on a large scale [23]. Drivers and other traffic participants can either appropriately trust, distrust or overtrust an automation system. Distrust occurs when humans' trust falls below a system's actual capabilities, whereas overtrust means that one excessively trusts automation even in situations the automation cannot handle. The goal of trust research is to "calibrate" users' subjective trust to a level where it matches a system's objective capabilities. Trust was widely addressed in studies on driver-vehicle interaction in the last years and remains an essential requirement in AV-human interaction [24, 29, 31, 68].

3 METHOD

The following section presents details regarding study planning and execution, as well as analysis methods.

3.1 Study Design

The study tested one independent variable, which was the type of communication cue on the AV in three different levels. The first condition was acceleration intention. This condition was demonstrated similarly to Mirnig et al. [47], with a white bar extending sideways repetitively on a LED matrix attached to the radiator grill of AV. The second condition, -deceleration intention- was demonstrated with two white bars moving and merging in the center of the matrix (Figure 2). Both of these designs were based on literature [47], expert opinions, and a short round of interviews with individuals, as well as field testing for visibility (Appendix A). Since peripheral vision is more specialized in detecting movement, we opted for animation patterns [25]. As closer objects are perceived as bigger and distant objects are smaller, we extended the light animation to imitate a growing and approaching object in acceleration intention while using a shrinking animation pattern in deceleration intention [6]. Lastly, in the baseline condition, the display did not show anything. Each condition was presented to each participant three times in a pseudo-randomized order.

Dependent variables were (1) binomial driving choices of participants as waited or passed first, (2) Situational Prosocial and Aggressive Behavior in Traffic Scale (SPAT) (in preparation), and Situational Trust Scale for Automated Driving (STS-AD) [32]. SPAT has 22 7-point semantic differential items made of adjectives in opposite poles. The middle point indicates neutral evaluation. Higher composite average scores indicate that the road user is evaluated as more prosocial by participants. STS-AD composes 6 items measuring situational trust score on a 7-point Likert scale form (1 = fully disagree, 7 = fully agree).

Furthermore, to be informed about the general sample profile, Prosocial and Aggressive driving Inventory (PADI) [28] and Prosocial Tendencies Measure Revised PTM-R [8, 61] were used. PADI has 29 statements which are constructed as 6-point Likert scale items, 1 indicating never acting as described in the statement, and 6 indicating always acting as described in the statement. PTM-R has 15 items with five points, 1 indicating stated behavior "does not describe me at all" and 5 indicating "describes me greatly".

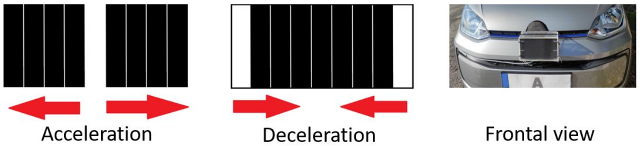

Figure 2: From left to right: Acceleration intention eHMI with a white bar extending sideways repetitively; deceleration intention eHMI with two white bars extending and merging in the center repetitively; frontal view of AV and eHMI attachment on radiator grill. In no eHMI condition the display was off.

3.2 Apparatus

A 64 x 32 flexible RGB LED matrix[1] with an Adafruit RGB matrix HAT [2] was used for eHMI. The matrix was programmed on a Raspberry Pi 4 Computer (Model B 2GB RAM) with Python Version 3.9. The connection to the Raspberry Pi was built with PuTTy [10], and Thonny [67] was used for programming with Python. The matrix was programmed with the rpi-rgb-led-matrix library [69]. To make the matrix work without an internet connection, a Python autostart script was written, which starts the LED-image-viewer directly after the Raspberry Pi is booted. To remotely switch between different displays, i.e. acceleration or deceleration, a simple PowerPoint Presenter with a USB receiver was used. The AV used in the study was a manually driven Volkswagen e-UP. The driver was hidden under a car seat costume similar to Rothenbücher et al. [62]. Stickers indicating automated driving were placed on the sides and the hood of the car (Figure 5). A dysfunctional Microsoft Xbox 360 Kinect was placed on the roof to simulate a sensor attachment. A custom-made LED matrix with a plexiglass casing was attached with a thin rope and cable binders to the radiator grills (Figure 2). Lastly, branding and license plate were covered.

3.3 Participants

24 participants (8 female, 16 male, age range 20 - 67, $M = 30.21$, $SD = 13.44$ years) took part in the study. Selection criteria were being over 18 years old, holding a driver's license ($M = 15.96$ years, $SD = 13.22$), owning a car, and having a normal or corrected-to-normal vision. Their average prosocial ($M = 4.9$ $SD = 0.45$) and aggressive driving scores ($M = 2.32$ $SD = 0.41$) indicated an overall positive and non-aggressive driving style. Their average composite PTM-R results ($M = 2.79$ $SD = 0.48$) signify neutral to a small prosocial tendency in the overall sample. Participation was compensated with 12 euros per hour and 30 cents travel costs per kilometer. They were reached online and with printed flyers. The ethics committee approved the study according to the Declaration of Helsinki.

3.4 Procedure

All participants were sent an online pre-questionnaire form that could be filled up voluntarily. On the experiment day, participants were invited to the test area with their own cars. AV was parked away from the reception area with the ghost driver inside. Upon arrival, the experimenter provided a consent form, study information and demographics document in the reception area. Afterward, the experimenter drove with the participant through the driving path and explained the tasks. (Figure 3). Locomotion intention eHMIs and their meanings were also introduced. In the meantime, the ghost driver drove to the starting position of AV. Then, the experimenter positioned herself near the narrow passage and started trials by counting until 3 over walkie-talkies, where both the participant and ghost driver could hear simultaneously (Figure 4). The ghost driver adapted her driving speed according to the participants' driving speed, in order to approach the narrow passage at the same time. Yet, the ghost driver left enough distance and time to enable the participant to make the decisive move to resolve the conflict. In line with Rettenmaier et al. [58], implicit locomotion cues of AV were matched with locomotion intention eHMI conditions. In other words, in acceleration eHMI condition, the ghost driver approached the narrow area more assertively with constant speed, while in deceleration eHMI condition she drove with a more defensive style. Lastly, in the neutral condition, a neutral driving style was adopted. The baseline speed was 10-12 km/h unless the participant was a very slow or fast driver. Consequently, the participant either slowed down and stopped, or continued driving and took priority to pass the narrow area. Both parts followed their paths and reached their starting points. Then, the experimenter asked how the driving or waiting decision was formed and to which aspects the participant paid attention. Afterwards, the participant filled out the intermediate questionnaires. After 9 repetitions, the participant drove back to the reception area and filled out post-questionnaires and answered to post-interview questions. On average, each trial including intermediate questionnaires took 3.5 minutes, while the entire study took 90 minutes per participant.

3.5 Analytical Approach

The interviews were analyzed employing an inductive category development [43, 66] to identify the frequency and distribution of specific words or phrases in the transcripts. We adopted this approach in exchange for Thematic Analysis, as our interviews did not include many statements with underlying feelings and emotions, but rather rich with recurrent words and phrases. Since each question had a specific theme such as usefulness or attention, they were treated as primary codes. Then, the answers were inspected, and similar themes were coded and summarized with code categories, which enabled statements with similar meanings to be grouped into joint code categories. For instance, attention was a predefined code since the question "What did you pay attention to?" would be giving attention-related answers. "AV reached the gap first" or "AV was too far" would be two different codes under attention-related answers, which eventually be merged under "distance". For the answers given after each trial, the number of occurrences of the same code over 72 trials was reported. Pre-questionnaire (N = 28) and post-questionnaire (N=24) answers were reported per participant. Participants could contribute to multiple codes if they answered with multiple themes. The codes were created by the second author and were discussed with the first author. All qualitative analysis steps were performed in the software MAXQDA Version 2022 [42].

Quantitative analysis steps were performed in RStudio (version 2023.03.0+386) [63]. To analyze the driving choices of participants, a

[1] https://www.adafruit.com/product/3826,[Online; accessed 15-March-2023]

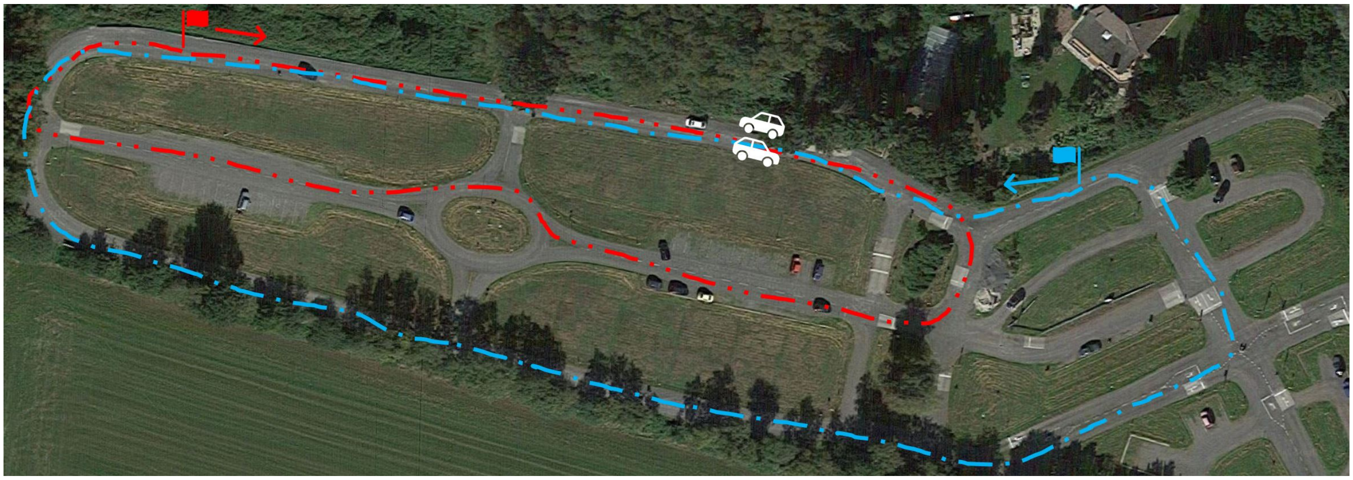

Figure 3: Driving paths of AV (blue) and participants (red). Flags indicate trial starting positions. White cars pinpoint the narrow passage where implicit negotiation happens.

generalized linear mixed-effects model (GLMM) [52] was calculated, by using the glmer function of the Lme4 package (version 1.1-27.1) [4]. Since the decisions of the participants were binomial, binomial family with logit link option to the regression model was added. The analyses of STS-AD and SPAT were done with two separate linear mixed-effects models (LMM), by using lmer function. In all models, eHMI conditions were added as fixed effects. Within-subject variance, sex, and age-related variability were added as random effects factors.

4 RESULTS

This section presents quantitative and qualitative results regarding experimental conditions only (acceleration and deceleration intention, and baseline eHMI). In the discussion section, further insights gained from qualitative results are shared.

4.1 Effects of Locomotion Intention on Driving Choices, SPAT, and STS-AD

Cumulative driving choices and average SPAT and STS-AD scores are given in Table 2. Indicating deceleration intention significantly increased waiting probabilities of participants (β = 2.34, $z(216)$ = 5.22, $Pr(>|z|) < .001$). Indicating acceleration intention significantly decreased waiting probabilities compared to the baseline level of no indication of locomotion intention (β = -0.81, $z(216)$ = -2.26, $Pr(>|z|) < .05$). Furthermore, the indication of deceleration intention significantly increased the probability of AV being perceived as prosocial (β = 9.87, $t(199.76)$ = 3.47, $Pr(>|t|) < .001$). Indication of acceleration intention did not predict an increase in prosocial perception (β = 5.23, $t(199.76)$ = 1.84, $Pr(>|t|)$ = .06). Finally, the indication of deceleration (β = 2.05, $t(190.03)$ = 1.81, $Pr(>|t|)$ = .07) or acceleration did not predict any changes in situational trust scores compared to the baseline condition (β = 1.61, $t(190.03)$ = 1.42, $Pr(>|t|)$ = .15).

Table 1: GLMM and LMM results of eHMI conditions on Driving Choices, Situational Social Perception (SPAT), and Situational Trust (STS-AD)

Predictors	M1 Odds Ratios	M2 Estimates	M3 Estimates
(Intercept)	0.72	37.29 ***	20.22 ***
Acceleration Intention	0.44 *	5.24	1.61
Deceleration Intention	10.41 ***	9.88 ***	2.06
Random Effects			
σ^2	3.29	290.68	46.29
τ_0 ID	0.02	0.00	0.01
τ_0 Age	0.03	31.15	3.13
τ_0 Sex	0.14	0.00	1.58
ICC	0.05		0.09
N ID	24	24	24
N Age	17	17	17
N Sex	2	2	2
Observations	216	216	216
Marginal r^2 / Conditional r^2	0.342 / 0.376	0.053 / NA	0.015 / 0.106

*** $p < .001$, ** $p < .01$, * $p < .05$

Note: *GLMM Odds ratios and Random Effects are reported for model 1. LMM Estimates and Random Effects are reported for model 2. M1: eHMI on driving choices, M2: eHMI on situational prosocial perception (SPAT), M3: eHMI on situational trust (STS-AD)*

Table 2: Driving choices and descriptive statistics of SPAT and STS-AD evaluations

	Acceleration	Deceleration	Baseline
Number of times Passed/Waited	19/53	64/8	32/40
Aver. Social Perception (SD)	5.21 (0.69)	5.43 (0.74)	4.9 (0.91)
Aver. Trust (SD)	5.39 (0.82)	5.44 (0.82)	5.14 (0.90)

4.2 Qualitative Feedback: The Effects of Locomotion Intention on Driving Choices

In 53 out of 72 trials (74%), participants decided to wait when faced with an acceleration intention eHMI. Of all the reasons given, half

Figure 4: Participant POV, driving towards narrow passage.

Figure 5: AV used in the study

of the answers (26/48) indicated that acceleration intention eHMI affected their decision to wait. Some others (7/24) stated that they did not pay attention to the display in every trial, with four of them only not paying attention to the display at the beginning and three others sometimes not paying attention to the display. Overall, some (7/48) of the answers included feelings, meaning that participants perceived the car was driving more assertively than usual. Other answers included lateral position (2/48), speed (2/48) of the car, and its distance (4/48) from parked cars.

Answers explaining the reason participants chose to drive first in the acceleration intention condition were diverse. In some cases (8/23), participants wanted to test the capabilities of the AV. In doing so, participants were pleased that the AV prioritized safety and waited until the participant safely passed the narrowing. However, not all of them wanted to test the AV, but they paid attention to the speed (9/23), and distance (2/23) of the AV and decided to take priority accordingly. They either felt reaching the narrow area faster or the AV hesitated to take priority. There were other reasons, such as S15, misinterpreting the display or two other participants already planning to take priority before the trial, regardless of the AV's intention.

With the deceleration intention display, 89% of the time (64/72) participants chose to pass first. In half of the reasons given (35/70), the deceleration intention display was the reason for deciding to take priority. In 5 trials the participants were insecure because the display did not harmonize with the intention, i.e. the AV drove too fast, although indicated to decelerate. Yet, despite the uncertainty, participants still chose to trust the display and passed. However, the display was not the only source for deciding to take priority. 22 out of 70 reasons given were related to AV's locomotion cues such as position, distance and especially speed. S19, for example, stated that she was paying attention to the display and acted accordingly, but she had nevertheless waited until the car stopped before passing. The reasons for participants not driving despite deceleration intention eHMI varied. Two participants intentionally tested the AV by giving priority to it, and they observed how the car reacted. Two participants did not drive in order to be cautious, since it was their first trial. One participant misinterpreted the eHMI sign and another participant felt that the display did not match the car's driving behavior, so they preferred to wait. For one participant, the AV was too close to parked cars, so she was unsure if her car could fit through, hence she let the AV drive.

For no eHMI (i.e., baseline condition), driving or waiting decisions were fairly balanced. Participants decided to pass in 56% of the trials (32/72) and waited in 44% of the trials (40/72). Participants stated paying the most attention to speed, lateral position, and distance (17/32 when waiting and 25/35 when passing). Thus, many of them chose to drive when the AV slowed down or stopped, and waited when they felt the AV would not stop. Another important decision factor was assessing who reached the narrowing first. Participants were more likely to let the AV pass if they perceived the AV reaching the parked cars first. Similarly, they insisted on their priority if they arrived first. Another aspect was the lateral position of AV. If the AV was driving in the middle of the road, the participants assumed that it would drive through. If it drove more on the right, the participants thought that it would wait. Some participants assumed that the AV would behave defensively if it did not have a display on, which is why they passed. Lastly, 8 out of 32 of the reasons for waiting and 3 out of 35 of the reasons for passing included statements regarding uncertainty because the display was off.

5 DISCUSSION

The following section elaborates on experimental results and brings up new insights gained from qualitative interview results regarding participants' expectations from AVs.

5.1 Explicit Locomotion Intention acts as a mediator for resolving traffic conflicts

Our results indicated a significant regulatory effect of acceleration and deceleration intention displays on participants' decision-making processes. Furthermore, participants stated explicitly that they took the information on the display into account while making their decision, together with actual locomotion cues and lateral positioning [45, 46, 58] of the AV on the road, which was also expressed in previous research [16, 65]. Overall, we could validate the findings of Rettenmaier and Bengler [57] from their driving simulation study in a realistic setting, that AV should indicate its intention implicitly and explicitly to solve conflicts in ambiguous scenarios easier.

5.2 Locomotion Intention Can Support Social-embeddedness of AVs

We found that when AV indicated deceleration intention with its display and with more defensive locomotion maneuvers, individuals' likelihood to perceive it as more prosocial increased compared to baseline (AV not indicating any intentions coupled with a neutral driving style). This emphasizes the preference of humans regarding the defensive behavior of AVs with clear intentions in ambiguous driving scenarios, which was further emphasized in the qualitative feedback our participants gave. This desire also collaborates with the findings of Miller et al. [46], where they run a similar bottleneck scenario with an AV on a driving simulator. Moreover, the perception of social behavior was indifferent when AV signaled acceleration intention or did not indicate any explicit intention with its display. Even though the clear acceleration intention may have made participants content, the AV asking for priority might have neutralized their perception and resulted in similar evaluations to the baseline condition. On the deceleration intention condition, however, we might have observed a highlighted effect of both aspects of clear indication and yielding behavior, which resulted in a significant prosocial perception of AV when compared to the baseline condition. Moreover, during the interviews, participants responded positively regarding the displays, and viewed them as helpful, in line with previous studies [11, 15]. Overall, these results indicate that, for AVs to be perceived as social, they do not have to equip anthropomorphic features [9, 39], and emphasized locomotion intention cues might partially substitute the missing validation cues of human-human interaction. The high focus on safety, predictability, and rule compliance in interview answers suggests that AVs are perceived as more positive and social if they had reliable functionality.

5.3 Situational Trust in Automation Requires a Wider Perspective

The average situational trust scores of our participants did not indicate any change in their trust across locomotion intention conditions. One potential reason could be that the items in STS-AD were primarily tailored for AV drivers, not for the other drivers interacting with an AV externally. Hence, we believe that the average of the entire item set may not have been sensitive enough to reveal changes in trust in different conditions in our experiment. Extending the perspective of the driver in STS-AD items (from inside to outside) could be a new way to utilize the scale for a wider range of situations. Furthermore, existing studies have mostly investigated trust in the context of eHMIs from the perspective of pedestrians, with mixed results. A study by Liu et al. [37] suggests that pedestrians trust AVs similar to manual vehicles in crossing situations. Similar results were obtained in an experiment by Bonneviot et al. [5], who showed that a communication HMI can increase pedestrians' subjective trust similar to encounters with human drivers. Since pedestrians are more vulnerable than drivers, and given the low speeds as present in our experiment, maybe the situations were not risky enough to require higher trust levels. Furthermore, in our experiment the eHMI status always matched the behavior of the AV, while in related experiments with pedestrians, trust often varied after experiencing automation failures [19, 30]. Future experiments may investigate this issue in the context of vehicle-to-vehicle interactions as well. Lastly, to get additional feedback in the post-interviews, participants were asked about their general trust in AV during the experiment. Some participants (9/24) trusted the AV from the beginning, considering that a safety-critical situation cannot occur during a controlled test. The other participants (9/24) reported that although they felt insecure, cautious, or unsafe at the beginning, they felt secure and confident as they proceeded and as they got used to encountering the AV.

5.4 Mixed Traffic with AVs: Expectations and Challenges

When participants were asked in pre-questionnaires regarding their expectations from AVs and how they envision future mixed traffic, the most important themes seemed to be rule compliance (23/28), predictable behavior (8/28), and safety (11/28). Furthermore, participants wanted AVs to behave thoughtfully (4/28), reliably (3/28), and defensively (3/28) [45]. Participants mentioned that AVs would be more efficient (7/28) in urgent situations such as forming an emergency lane (3/28). However, concerns and negative effects were also raised. For example, P05 was concerned that the car might fail to detect people and thus increase the accident rate. Others described AVs as untrustworthy, confusing, and uncertain. P04 stated the reason as the fear of giving up on control, which makes individuals feel insecure. P05 justified his fear by stating his worry of AV overlooking them. The general opinion seemed to be that AVs should be subordinate to humans and not insist on their right of way. In this way, accidents could be avoided and safety could be ensured. Moreover, AVs are expected to strictly follow traffic rules. However, strict adherence to traffic rules might inhibit AVs to react to ambiguous situations, especially caused by rule-braking human drivers.

When participants were asked how AVs should react in ambiguous situations where the right of way is not clear in post-interviews, participants proposed either a defensive (11/24) or reactive (13/24) approach. Participants who argued for defensive behavior stated that an AV should always be passive in potential conflicts and give the right of way to humans, similar to the direction of answers reported by Miller et al. [45]. Some of the participants who advocated strict adherence to traffic rules nevertheless mentioned a few exceptions such as safety-critical situations where deviation from rules would be necessary. Participants arguing for a reactive approach exposed their wish to have communication with AV, through signaling with flashing lights or displays. They further expressed that AVs should adapt to humans by analyzing their behavior, speed, and position to decide whether they should wait or take priority, as strict enforcement of driving rules would otherwise be too unfamiliar or annoying. Two participants specifically stated that AVs should have defensive behavior, while P09 argued that AVs should not only follow traffic rules but also insist on their rights, as this would make them less likely to be taken advantage of. Three participants argued that a distinction should be made between traffic with only autonomous vehicles and mixed traffic, as humans can disrupt the traffic flow through unexpected behavior. Thus, it was suggested that in pure traffic with only AVs, vehicles should be strictly rule-abiding, while in mixed traffic, AVs should act more human-like and sometimes deviate from the rules to resolve conflicts. Two other participants suggested special lanes for AVs similar to bus lanes, where they could have additional regulations while still being a part of urban traffic.

During the experiment, some participants wanted to test the capabilities of the AV and forced the ghost driver to make a decision. In one instance, both AV and the participant kept driving very slowly to see who breaks the ambiguity and make the decision first. The other two times ghost driver had to reverse since she was already in the narrow area, yet the participants wanted to see what would happen if they change their minds and insist on taking priority. Other times they chose to drive on even if AV indicated acceleration, to see if it was reactive. These instances emphasize similar playful reactions of individuals in Moore et al. [48]. It seems that when encountering a new technology, some individuals will likely act unexpectedly and potentially undermine the usability of AVs.

5.5 Methodological Implications, Limitations and Future Directions

We conducted a realistic field experiment on a bottleneck scenario naturally formed by two parked cars. Our participants drove their own cars or their family cars so that they would feel the most comfortable while driving. This aspect gave us the most naturalistic results we could obtain from a controlled test track study. We did not reveal to our participants that AV was manually driven until the end of the experiment. This resulted in a successful manipulation of the perception regarding the automation status of the car, which we validated with post interviews. None of our participants suspected that AV was manually driven and only a few considered the possibility of AV being remotely controlled, or someone might be sitting in the back seat. Conducting this study on a test track in a realistic condition extends the existing research with results that have higher external validity in terms of perception of auditory, visual, and vestibular/motion cues [33]. Furthermore, the perception of risks in field studies is generally higher than in driving simulators [17]. However, it is important to point out several limitations. Due to the lack of access to AVs, and the safety-critical nature of the tested scenarios, it was ethically impossible to conduct such an experiment in an uncontrolled environment. Nevertheless, as our study mainly explored the effect of an eHMI in such a scenario , it still maintains a high relative validity of the effects observed in comparison to previous studies conducted in driving simulators. Furthermore, the controlled design of the study ensures internal validity of the effects observed [33]. As this study –to our knowledge– is the first one testing the eHMIs in real-life bottleneck scenarios, further research is required to understand whether the presence of other variables will have an effect on the results observed. Moreover, The US National Highway Traffic Safety Administration (NHTSA) Standing General Order on Crash Reporting (SGO)[2] indicate that the most crashes occur with passenger cars. Although we set up our study with passenger cars, there are wide range of other scenarios to be investigated beyond bottleneck situations.

Regarding the communication interface, we utilized an LED matrix where many signals could be easily shown. After pilot tests, we noted that the most visible cues were vertical bars in white. Our participants reported in the interviews that the display was visible (19/24). Most of them understood the communication cue on the display with ease (19/24). S07 found the acceleration display more understandable than the deceleration display, which is why he suggested only using an acceleration indicator. Almost all the participants found the display helpful (23/24), because it helped to compensate for the missing driver-driver communication, and provided additional feedback and guidance. Only S23 found the display more annoying and distracting than helpful. Data collection took an entire month, under adverse and good weather conditions including fog, snow storms, strong winds, heavy rain, and bright sun. The display was robust and visible through all conditions. On sunny days when the sunlight directly reflected on the display the visibility decreased, however, it was still sufficient at closer distances. Under the snow storm, we had to clean the snow on the eHMI after every trial. We did not run the experiments at sunset, as the visibility under the car seat costume decreased significantly.

Participants were asked if they had any suggestions for improvement. Five people wished that AV would indicate their intention with flashing lights. Some participants suggested using blinkers since the AV didn't use blinkers in the experiment to limit confounding variables. Two participants wished that the AV would brake faster or keep more distance from the narrowing. For the improvements regarding the communication interface, some participants wished for colors (11/24), in particular red and green. There was no consensus regarding the meaning of the colors among participants. Some suggested that it should be about AVs' intention, while others suggested that it should indicate what others should do as in traffic lights, which further validates the frame of reference problem in the use of traffic light colors [14]. We were able to prevent different perceptions of the cues by introducing the interface before the experiment, yet in the wild, introduction of abstract cues may create confusion among road users. This calls for further realistic and longitudinal studies in order to grasp the long term impact of eHMIs on road users. In our study, 90 minute exposure could only serve as an introduction to these novel interfaces. Participants further wanted the display to be larger (7/24), and indicate other signs including vertical arrows, a circular sun, a traffic light, and a human face (7/24). S22 liked the idea of displaying human needs on a display, such as a pregnant woman or being late. Last but not least, four participants suggested changing the position of the display, ideally more around the eye level. S13 suggested digitizing the license plate as well and alternating it with the display whenever it would be needed.

Our eHMIs signaled locomotion intention explicitly, however, they were not dynamic. This means that they were not adaptive

[2]https://www.nhtsa.gov/laws-regulations/standing-general-order-crash-reporting#ads,[Online; accessed 19-June-2023]

to the actual speed changes of the car. Participants encountered one design per condition and the eHMI they saw in each condition strictly indicated the intention of the AV in the bottleneck situation. Yet, the display was on from the beginning, until the end of the trial. The display could have been more adaptive and situation-specific, where it could be turned off once the ambiguity was resolved. Furthermore, we had only 24 participants due to short-notice dropouts. However, their backgrounds, ages, and driving experience were diverse. As the next steps, not only driving decisions but also actual driving behavior of participants and driving behavior of ghost driver could be analyzed and another perspective on driver behavior could be presented. Furthermore, quantitative validation of where participants paid attention during the encounter could be achieved with mobile eye trackers. Lastly, STS-AD can be adapted for measuring the situational trust of drivers outside the AV.

6 CONCLUSION

We conducted a realistic test track study in an ambiguous bottleneck scenario where participants drove their own cars. They had to decide whether to give or yield the crossing priority to a self-driving vehicle, which was manually driven by a human under a car seat costume in reality. AV was equipped with an external display indicating the locomotion intention of the car explicitly with vertical bars. We used mixed methods to evaluate the effects of yielding intention and acceleration intention compared to the baseline condition. Our results revealed that deceleration intention significantly decreased yielding probabilities and significantly increased prosocial perception of the AV, and it was desired as the default behavior of AV in ambiguous scenarios by our participants. Furthermore, acceleration intention significantly increased yielding probabilities but did not have any effect on prosocial perception compared to the baseline condition. Situational trust was not affected by different conditions in our study. Participants reported using explicit locomotion intention cues together with implicit locomotion cues such as lateral movement, speed, and distance of the AV. Lastly, they found explicit locomotion intention cues, understandable, visible, and helpful.

ACKNOWLEDGMENTS

We thank Dr. Heiko Müller, Dr. Marion Koelle, Susanna Krämer, Tobias Lunte, Jonas Keil, Jonathan Kutter, Bastian Brandes and Human-Centered AI Group members at OFFIS for their valuable contributions.

REFERENCES

[1] Sander Ackermans, Debargha Dey, Peter Ruijten, Raymond H Cuijpers, and Bastian Pfleging. 2020. The Effects of Explicit Intention Communication, Conspicuous Sensors, and Pedestrian Attitude in Interactions with Automated Vehicles. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM SIGCHI, Hawaii, USA, 1–14.

[2] Adafruit. 2023. Adafruit RGB Matrix HAT + RTC for Raspberry Pi - Mini Kit. https://www.adafruit.com/product/2345,[Online; accessed 15-March-2023].

[3] Dina AlAdawy, Michael Glazer, Jack Terwilliger, Henri Schmidt, Josh Domeyer, Bruce Mehler, Bryan Reimer, and Lex Fridman. 2019. Eye Contact between Pedestrians and Drivers. In *Driving Assessment Conference*. The University of Iowa, Iowa, USA, 301–307. https://doi.org/10.17077/drivingassessment.1710 arXiv:1904.04188

[4] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1 (2015), 1–48. https://doi.org/10.18637/jss.v067.i01

[5] Flavie Bonneviot, Stéphanie Coeugnet, and Eric Brangier. 2021. Pedestrians-Automated Vehicles Interaction: Toward a Specific Trust Model?. In *Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021) Volume III: Sector Based Ergonomics*. Springer, Springer, Virtual, 568–574.

[6] Shadan Sadeghian Borojeni, Lars Weber, Wilko Heuten, and Susanne Boll. 2018. From reading to driving: priming mobile users for take-over situations in highly automated driving. In *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services*. ACM SIGCHI, Barcelona, Spain, 1–12.

[7] Fanta Camara, Serhan Cosar, Nicola Bellotto, Natasha Merat, Charles Fox, et al. 2020. Continuous Game Theory Pedestrian Modelling Method for Autonomous Vehicles. , 20 pages.

[8] Gustavo Carlo, Anne Hausmann, Stacie Christiansen, and Brandy A Randall. 2003. Sociocognitive and behavioral correlates of a measure of prosocial tendencies for adolescents. *The journal of early adolescence* 23, 1 (2003), 107–134.

[9] Chia-Ming Chang, Koki Toda, Daisuke Sakamoto, and Takeo Igarashi. 2017. Eyes on a Car: an Interface Design for Communication between an Autonomous Car and a Pedestrian. In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications*. ACM SIGCHI, Oldenburg, Germany, 65–73.

[10] chiark.greenend. 2023. PuTTY: a free SSH and Telnet client. https://www.chiark.greenend.org.uk/~sgtatham/putty/,[Online; accessed 15-March-2023].

[11] Mark Colley, Tim Fabian, and Enrico Rukzio. 2022. Investigating the Effects of External Communication and Automation Behavior on Manual Drivers at Intersections. *Proceedings of the ACM on Human-Computer Interaction* 6, MHCI (Sept. 2022), 1–16. https://doi.org/10.1145/3546711

[12] Koen De Clercq, Andre Dietrich, Juan Pablo Núñez Velasco, Joost De Winter, and Riender Happee. 2019. External Human-Machine Interfaces on Automated Vehicles: Effects on Pedestrian Crossing Decisions. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 61, 8 (2019), 1353–1370.

[13] Debargha Dey, Azra Habibovic, Andreas Löcken, Philipp Wintersberger, Bastian Pfleging, Andreas Riener, Marieke Martens, and Jacques Terken. 2020. Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transportation Research Interdisciplinary Perspectives* 7 (2020), 100174.

[14] Debargha Dey, Azra Habibovic, Bastian Pfleging, Marieke Martens, and Jacques Terken. 2020. Color and animation preferences for a light band eHMI in interactions between automated vehicles and pedestrians. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM SIGCHI, Hawaii, USA, 1–13.

[15] Debargha Dey, Andrii Matviienko, 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. *Information Technology* Submitted, Special Issue: Automotive User Interfaces in the Age of Automation (2020), 123–141. https://doi.org/10.1515/ITIT-2020-0025

[16] Debargha Dey and Jacques Terken. 2017. Pedestrian interaction with vehicles: roles of explicit and implicit communication. In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications*. ACM SIGCHI, Oldenburg, Germany, 109–113.

[17] Leonard Evans. 1991. *Traffic safety and the driver.* Science Serving Society, MI, USA.

[18] Roja Ezzati Amini, Christos Katrakazas, and Constantinos Antoniou. 2019. Negotiation and decision-making for a pedestrian roadway crossing: A literature review. *Sustainability* 11, 23 (2019), 6713.

[19] Stefanie M. Faas, Johannes Kraus, Alexander Schoenhals, and Martin Baumann. 2021. Calibrating Pedestrians' Trust in Automated Vehicles. , 17 pages. https://doi.org/10.1145/3411764.3445738

[20] Stefanie M 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 (2020), 171–186.

[21] Charles Fox, Fanta Camara, Gustav Markkula, Richard Romano, Ruth Madigan, Natasha Merat, et al. 2018. When should the chicken cross the road?: Game theory for autonomous vehicle-human interactions.

[22] Lex Fridman, Bruce Mehler, Lei Xia, Yangyang Yang, Laura Yvonne Facusse, and Bryan Reimer. 2019. To Walk or Not to Walk: Crowdsourced Assessment of External Vehicle-to-Pedestrian Displays. arXiv:1707.02698 https://arxiv.org/pdf/1707.02698.pdfhttp://arxiv.org/abs/1707.02698

[23] Anna-Katharina Frison, Philipp Wintersberger, Andreas Riener, Clemens Schartmüller, Linda Ng Boyle, Erika Miller, and Klemens Weigl. 2019. In UX we trust: Investigation of aesthetics and usability of driver-vehicle interfaces and their impact on the perception of automated driving. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. ACM SIGCHI, Glasgow, Scotland, 1–13.

[24] Christian Gold, Moritz Körber, Christoph Hohenberger, David Lechner, and Klaus Bengler. 2015. Trust in automation–before and after the experience of take-over scenarios in a highly automated vehicle. *Procedia Manufacturing* 3 (2015), 3025–3032.

[25] E Bruce Goldstein and Laura Cacciamani. 2021. *Sensation and perception*. Cengage Learning, Boston, Massachusetts, United States.
[26] Nicolas Guéguen, Sébastien Meineri, and Chloé Eyssartier. 2015. A pedestrian's stare and drivers' stopping behavior: A field experiment at the pedestrian crossing. *Safety science* 75 (2015), 87–89.
[27] Azra Habibovic, Victor Malmsten Lundgren, Jonas Andersson, Maria Klingegård, Tobias Lagström, Anna Sirkka, Johan Fagerlönn, Claes Edgren, Rikard Fredriksson, Stas Krupenia, et al. 2018. Communicating intent of automated vehicles to pedestrians. *Frontiers in psychology* 9 (2018), 1336.
[28] Paul B Harris, John M Houston, Jose A Vazquez, Janan A Smither, Amanda Harms, Jeffrey A Dahlke, and Daniel A Sachau. 2014. The Prosocial and Aggressive Driving Inventory (PADI): A self-report measure of safe and unsafe driving behaviors. *Accident Analysis & Prevention* 72 (2014), 1–8.
[29] Sebastian Hergeth, Lutz Lorenz, Roman Vilimek, and Josef F Krems. 2016. Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving. *Human factors* 58, 3 (2016), 509–519.
[30] Kai Holländer, Philipp Wintersberger, and Andreas Butz. 2019. Overtrust in external cues of automated vehicles: an experimental investigation. In *Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications*. ACM SIGCHI, Utrecht, Netherlands, 211–221.
[31] Brittany E Holthausen, Philipp Wintersberger, Zoe Becerra, Alexander G Mirnig, Alexander Kunze, and Bruce N Walker. 2019. Third workshop on trust in automation: how does trust influence interaction. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*. ACM SIGCHI, Utrecht, Netherlands, 13–18.
[32] Brittany E Holthausen, Philipp Wintersberger, Bruce N Walker, and Andreas Riener. 2020. Situational trust scale for automated driving (STS-AD): Development and initial validation. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM SIGCHI, Virtual, DC, USA, 40–47.
[33] Nico A Kaptein, Jan Theeuwes, and Richard Van Der Horst. 1996. Driving simulator validity: Some considerations. *Transportation research record* 1550, 1 (1996), 30–36.
[34] Mirjam Lanzer, Franziska Babel, Fei Yan, Bihan Zhang, Fang You, Jianmin Wang, and Martin Baumann. 2020. Designing communication strategies of autonomous vehicles with pedestrians: an intercultural study. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM SIGCHI, Virtual, DC, USA, 122–131.
[35] John D Lee and Katrina A See. 2004. Trust in automation: Designing for appropriate reliance. *Human factors* 46, 1 (2004), 50–80.
[36] Yee Mun Lee, Ruth Madigan, Oscar Giles, Laura Garach-Morcillo, Gustav Markkula, Charles Fox, Fanta Camara, Markus Rothmueller, Signe Alexandra Vendelbo-Larsen, Pernille Holm Rasmussen, Andre Dietrich, Dimitris Nathanael, Villy Portouli, Anna Schieben, and Natasha Merat. 2020. Road users rarely use explicit communication when interacting in today's traffic: implications for automated vehicles. *Cognition, Technology and Work* 1 (2020), 3. https://doi.org/10.1007/s10111-020-00635-y
[37] Hailong Liu, Takatsugu Hirayama, and Masaya Watanabe. 2021. Importance of instruction for pedestrian-automated driving vehicle interaction with an external human machine interface: Effects on pedestrians' situation awareness, trust, perceived risks and decision making. In *2021 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, Institute of Electrical and Electronics Engineers (IEEE), Nagoya, Japan, 748–754.
[38] Peng Liu, Yong Du, Lin Wang, and Ju Da Young. 2020. Ready to bully automated vehicles on public roads? *Accident Analysis & Prevention* 137 (2020), 105457.
[39] Andreas Löcken, Carmen Golling, and Andreas Riener. 2019. How should automated vehicles interact with pedestrians? A comparative analysis of interaction concepts in virtual reality. In *Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications*. ACM SIGCHI, Utrecht, Netherlands, 262–274.
[40] Victor Malmsten Lundgren, Azra Habibovic, Jonas Andersson, Tobias Lagström, Maria Nilsson, Anna Sirkka, Johan Fagerlönn, Rikard Fredriksson, Claes Edgren, Stas Krupenia, et al. 2017. Will there be new communication needs when introducing automated vehicles to the urban context? In *Advances in human aspects of transportation*. Springer, NY, USA, 485–497.
[41] Gustav Markkula, Ruth Madigan, Dimitris Nathanael, Evangelia Portouli, Yee M Lee, André Dietrich, Jac Billington, Anna Schieben, and Natasha Merat. 2020. Defining interactions: A conceptual framework for understanding interactive behaviour in human and automated road traffic. *Theoretical Issues in Ergonomics Science* 21, 6 (2020), 1–24.
[42] MAXQDA. 2023. The #1 software for qualitative and mixed methods data analysis. https://www.maxqda.com/,[Online; accessed 17-March-2023].
[43] Philipp Mayring et al. 2004. Qualitative content analysis. *A companion to qualitative research* 1, 2 (2004), 159–176.
[44] Adam Millard-Ball. 2018. Pedestrians, autonomous vehicles, and cities. *Journal of planning education and research* 38, 1 (2018), 6–12.
[45] Linda Miller, Ina Marie Koniakowsky, Johannes Kraus, and Martin Baumann. 2022. The Impact of Expectations about Automated and Manual Vehicles on Drivers' Behavior: Insights from a Mixed Traffic Driving Simulator Study. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, Seoul Republic of Korea, 150–161. https://doi.org/10.1145/3543174.3546837
[46] Linda Miller, Jasmin Leitner, Johannes Kraus, and Martin Baumann. 2022. Implicit intention communication as a design opportunity for automated vehicles: Understanding drivers' interpretation of vehicle trajectory at narrow passages. *Accident Analysis & Prevention* 173 (Aug. 2022), 106691. https://doi.org/10.1016/j.aap.2022.106691
[47] Alexander G Mirnig, Magdalena Gärtner, Peter Fröhlich, Vivien Wallner, Anna Sjörs Dahlman, Anna Anund, Petr Pokorny, Marjan Hagenzieker, Torkel Bjørnskau, Ole Aasvik, et al. 2022. External communication of automated shuttles: Results, experiences, and lessons learned from three European long-term research projects. *Frontiers in Robotics and AI* 9 (2022), 239.
[48] Dylan Moore, Rebecca Currano, Michael Shanks, and David Sirkin. 2020. Defense against the dark cars: Design principles for griefing of autonomous vehicles. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. ACM / IEEE, Cambridge, UK, 201–209.
[49] Dylan Moore, Rebecca Currano, G Ella Strack, and David Sirkin. 2019. The case for implicit external human-machine interfaces for autonomous vehicles. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM SIGCHI, Utrecht, Netherlands, 295–307.
[50] Lars Müller, Malte Risto, and Colleen Emmenegger. 2016. The social behavior of autonomous vehicles. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM SIGCHI, Heidelberg, Germany, 686–689.
[51] Cassidy Myers, Thomas Zane, Ron Van Houten, and Vincent T Francisco. 2022. The effects of pedestrian gestures on driver yielding at crosswalks: A systematic replication. *Journal of applied behavior analysis* 55, 2 (2022), 572–583.
[52] John Ashworth Nelder and Robert WM Wedderburn. 1972. Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)* 135, 3 (1972), 370–384.
[53] Anatol Rapoport and Albert M Chammah. 1966. The game of chicken. *American Behavioral Scientist* 10, 3 (1966), 10–28.
[54] Amir Rasouli, Iuliia Kotseruba, and John K Tsotsos. 2017. Agreeing to cross: How drivers and pedestrians communicate. In *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, Institute of Electrical and Electronics Engineers (IEEE), Redondo Beach, California, USA, 264–269.
[55] Michael Rettenmaier, Deike Albers, and Klaus Bengler. 2020. After you?! – Use of external human-machine interfaces in road bottleneck scenarios. *Transportation Research Part F: Traffic Psychology and Behaviour* 70 (April 2020), 175–190. https://doi.org/10.1016/j.trf.2020.03.004
[56] Michael Rettenmaier and Klaus Bengler. 2020. Modeling the Interaction with Automated Vehicles in Road Bottleneck Scenarios. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64, 1 (Dec. 2020), 1615–1619. https://doi.org/10.1177/1071181320641391
[57] Michael Rettenmaier and Klaus Bengler. 2021. The Matter of How and When: Comparing Explicit and Implicit Communication Strategies of Automated Vehicles in Bottleneck Scenarios. *IEEE Open Journal of Intelligent Transportation Systems* 2 (2021), 282–293. https://doi.org/10.1109/OJITS.2021.3107678
[58] Michael Rettenmaier, Sabrina Dinkel, and Klaus Bengler. 2021. Communication via motion – Suitability of automated vehicle movements to negotiate the right of way in road bottleneck scenarios. *Applied Ergonomics* 95 (Sept. 2021), 103438. https://doi.org/10.1016/j.apergo.2021.103438
[59] Michael Rettenmaier, Moritz Pietsch, Jonas Schmidtler, and Klaus Bengler. 2019. Passing through the Bottleneck - The Potential of External Human-Machine Interfaces. In *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, Paris, France, 1687–1692. https://doi.org/10.1109/IVS.2019.8814082
[60] Malte Risto, Colleen Emmenegger, Erik Vinkhuyzen, Melissa Cefkin, and Jim Hollan. 2017. Human-vehicle interfaces: the power of vehicle movement gestures in human road user coordination.
[61] Johannes Rodrigues, Natalie Ulrich, Patrick Mussel, Gustavo Carlo, and Johannes Hewig. 2017. Measuring prosocial tendencies in Germany: Sources of validity and reliablity of the revised prosocial tendency measure. *Frontiers in psychology* 8 (2017), 2119.
[62] Dirk Rothenbücher, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. 2015. Ghost driver: a platform for investigating interactions between pedestrians and driverless vehicles. In *Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM SIGCHI, Nottingham, UK, 44–49.
[63] RStudio Team. 2020. *RStudio: Integrated Development Environment for R*. RStudio, PBC., Boston, MA. http://www.rstudio.com/
[64] Shadan Sadeghian, Marc Hassenzahl, and Kai Eckoldt. 2020. An exploration of prosocial aspects of communication cues between automated vehicles and pedestrians. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM SIGCHI, Virtual, DC, USA, 205–211.

[65] Matúš Šucha, Daniel Dostal, and Ralf Risser. 2017. Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention* 102 (2017), 41–50.
[66] David R Thomas. 2006. A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation* 27, 2 (2006), 237–246.
[67] Thonny. 2023. Thonny: Python IDE for beginners. https://thonny.org/,[Online; accessed 15-March-2023].
[68] Philipp Wintersberger, Clemens Schartmüller, Shadan Sadeghian, Anna-Katharina Frison, and Andreas Riener. 2021. Evaluation of imminent take-over requests with real automation on a test track. , 00187208211051435 pages.
[69] Henner Zeller. 2023. Controlling RGB LED display with Raspberry Pi GPIO. https://github.com/hzeller/rpi-rgb-led-matrix,[Online; accessed 15-March-2023].
[70] Xiangling Zhuang and Changxu Wu. 2014. Pedestrian gestures increase driver yielding at uncontrolled mid-block road crossings. *Accident Analysis & Prevention* 70 (2014), 235–244.

A NON-USED EHMI DESIGNS.

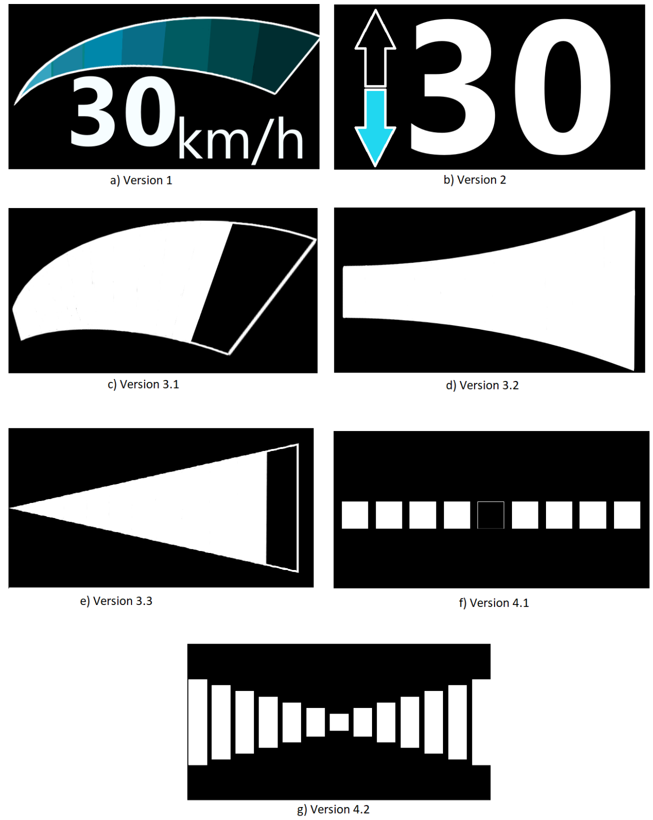

Figure 6: Some unused results of iterative design process of eHMIs. Blue color did have have ideal visibility in the field. Digits and animation combinations were found too crowded by pilot testers. Cone shaped bars were interpreted as turn indicators. Thin light bars had less visibility compared to full screen use of long light bars we eventually selected for our study.