Scalability in External Communication of Automated Vehicles: Evaluation and Recommendations

MARK COLLEY, Institute of Media Informatics, Ulm University, Germany
JULIAN BRITTEN, Institute of Media Informatics, Ulm University, Germany
ENRICO RUKZIO, Institute of Media Informatics, Ulm University, Germany

Automated vehicles will alter traffic fundamentally. While users can engage in non-driving-related tasks such as reading or even sleeping, the possibility to interact with other road users such as pedestrians via, for example, eye contact vanishes. Therefore, external communication of automated vehicles is currently researched with various concepts spanning dimensions such as anthropomorphism, technology, viewpoint, locus, message type, and others. However, the proposed concepts are mostly evaluated in simple scenarios, such as one person trying to cross in front of one automated vehicle. Therefore, we implemented a WebGL application of a four-lane road and conducted a within-subject study (N=46) to study the effects of nine concepts with and without the presence of other pedestrians and altering the yielding target of the automated vehicle. We found that all concepts were rated better than having no external communication. However, the effects were not uniform across the concepts.

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

Additional Key Words and Phrases: Automated vehicles; scalability; external communication; interface design.

ACM Reference Format:
Mark Colley, Julian Britten, and Enrico Rukzio. 2023. Scalability in External Communication of Automated Vehicles: Evaluation and Recommendations. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 2, Article 51 (June 2023), 26 pages. https://doi.org/10.1145/3596248

1 INTRODUCTION

Increased safety and facilitated mobility are some of the anticipated benefits of automated vehicles (AVs) [33]. AVs, as defined in SAE Level 5 [67], will be able to drive without any human passenger present. Otherwise, passengers could engage in a variety of non-driving-related activities, such as reading, sleeping, or working [41, 61]. Therefore, interpersonal communication, such as hand gestures and eye contact between drivers and pedestrians, will vanish. Current research, therefore, evaluates the necessity and the appropriateness of AVs substituting or even enhancing this communication process (e.g., [10]). While technological challenges have to be overcome, the acceptance of AVs will, ultimately, also depend on their safe integration and interaction with vulnerable road users in various scenarios. However, most of these investigations are based on simplistic one-on-one interactions (i.e., one AV communicates with one pedestrian) [18]. The aspect of being appropriate for a varying number of communication partners, such as bicyclists or pedestrians, and a varying number of vehicles, is called *scalability* [18]. While scalability (which could also be called situational complexity) is mentioned as a potential problem in some of these publications [18, 26, 52, 56, 57, 66], research has not yet addressed this issue specifically [18]. Scalability in this context addresses the ability of the external communication concept to be used in scenarios with varying numbers

---

Authors' addresses: Mark Colley, mark.colley@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany; Julian Britten, julian.britten@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany; Enrico Rukzio, enrico.rukzio@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany.

2474-9567/2023/6-ART51
https://doi.org/10.1145/3596248

of vehicles and/or pedestrians. Therefore, we have implemented a simulation with four lanes and numerous AVs and also included other pedestrians. Based on the approach of the *game with a purpose* [70] *The Walking Data* [44], we deployed this simulation onto a webpage[1]. With this, we conducted a user study (*N*=46). Participants used the classical WASD control and their mouse to move their gaze. In 36 trials, participants encountered **nine** external Human-Machine Interfaces (eHMIs). They encountered these once as the only participant (scenario *solo*) with the AV stopping for the participant and three times with two other pedestrians standing left and right at a distance of about 5m. In these three crossings, the AVs stopped for the left or the right pedestrian or the participant. AV behavior that does not fit to the crossing intention of pedestrians could occur if the AV falsely predicts this crossing intention. State-of-the-art approaches to pedestrian intention prediction still only achieve around 63% accuracy [73]. We compared the scenario *solo* with the condition with multiple pedestrians where the AV stopped for the participant. Also, we analyzed the effect of stopping behavior by evaluating the three conditions with multiple pedestrians.

We found that eHMIs, in general, are preferred to no eHMI. Additionally, eHMIs increase trust and reduce mental workload. There was a clear preference for the *Pedestrian Symbol*, the *Street Projection*, and the *Bumper Light*. In 1656 crossings recorded, three dangerous accidents occurred.

Contribution Statement: (1) Implementation of a WebGL game to study scalability aspects of eHMIs with **nine** eHMI concepts from related work included and adjusted and simulated surrounding pedestrians. (2) Findings of an online within-subjects study with *N*=46.

2 EXTERNAL COMMUNICATION OF AUTOMATED VEHICLES

In manual traffic, unclear traffic situations and traffic-related problems are overcome via gestures and eye contact [62]. Lee et al. [49] found that in Athens, Munich, and Leeds, the necessity for such gesture communication by the pedestrian (≈ 4%) and the driver (between 2 and 10%) was low. There was no data available on eye contact, but pedestrians looked towards the vehicle in the high majority of encounters [49]. Despite this low frequency, the external communication of AVs with vulnerable road users [42], also called external Human-Machine-Interfaces (eHMIs), was shown to be beneficial both at first encounters [13, 15] and in longitudinal studies [31].

Prior work focused on children [6, 20], people with vision [15, 16], mobility [3], or cognitive impairments [38], general pedestrians [2, 26, 52], and bicyclists [45]. Different modalities such as displays [7, 35], LED strips [23, 35, 54, 58] (and windshield displays [24]), movement patterns [76], projections [1, 60], external devices such as smartphones [43], auditory or tactile cues [57, 59] and combinations [57] or enhanced infrastructure [65] were proposed. Mostly, positive effects were found. For example, Faas et al. [31] found that perceived safety and crossing onset time improved with an eHMI, and the effect became stronger over time.

Various dimensions relate to the eHMI concept. These concepts can be grouped by modality, message type, and communication location [14] and more [22]. The *communication location* defines where the communication occurs: on the vehicle, the personal device, or the infrastructure (e.g., sidewalk). In this work, we solely explore communication that occurs directly on the AV. The position on the AV can be further broken down (top of the vehicle, bumper, windshield, on the sides, projection). This is relevant as Dey et al. [28] demonstrated that a pedestrian's eye-gaze shifts from the bumper to the windshield as the vehicle approaches. Additionally to the concept dimensions, Colley and Rukzio [14] describe situation parameters such as communication relationship (one-to-one, one-to-many, many-to-one, and many-to-many), acoustic noise, or communication partner (e.g., pedestrian or cyclist).

Most works focused on one-to-one scenarios [18], that is, one AV communicates with one pedestrian. Therefore, Colley et al. [18] described that the so-called "scalability" problem of AVs is not yet solved. With special regard to scalability, the recently proposed concept of Dey et al. [24] evaluated an eHMI with distance-dependent

---

[1] https://cross-the-road.onrender.com/

information (i.e., indicating when the AV will halt). This information was conveyed via a windshield that, from the middle to the side, was “filled” with cyan the closer the AV came to the relevant pedestrian. The results showed that this improved pedestrians’ comprehension of the AV’s intention and increased their willingness to cross.

A VR study conducted by Dietrich et al. [29] showed the problem of ambiguity of undirected eHMI concepts. In one scenario, a study participant was asked to cross the road at a zebra crossing. Another virtual pedestrian was crossing the street from the opposite side. A vehicle was approaching from the left, but the participant’s view was obstructed by a larger vehicle parking in front of the zebra crossing. The scenario was designed for the vehicle to yield to the virtual pedestrian and give them either a directed or an undirected signal to cross. If the vehicle recognized the participant, the light would either stay on (undirected) or turn towards them (directed). If the vehicle did not recognize the participant after the virtual pedestrian crossed, the light would turn off (undirected) or never be visible to them (directed; since the vehicle only targeted the virtual pedestrian). The study results show that if the AV used undirected light signals and only communicated with the virtual pedestrian, only 9 out of 30 participants understood the situation correctly.

Kaleefathullah et al. [46] conducted a study (*N*=60, 50 trials per participant) using a Cave Automatic Virtual Environment (CAVE) simulator to investigate trust development and potential misuse of eHMIs by pedestrians. The authors varied the yielding behavior of the AV (yielding at 43m or 33m distance or not yielding at all). The AV featured either no eHMI or a LED eHMI (light strips around the windshield and front grill). The study was split into two groups: For the first group, the eHMI would turn on 1 second before, and for the second group 1 second after beginning to yield. 2 failure trials were included where the AV’s eHMI would turn on, but the vehicle would drive at a constant speed and would not yield. Kaleefathullah et al. [46] found that people in the first group (activates early) and participants with an eHMI entered the road earlier. 35% of the participants entered the road in the first failure trial (see overtrust). In the group where the eHMI would activate 1 second before the vehicle started to slow down, the trust stayed lower, and the participants entered the road later after the first failure trial in comparison to before. During the second failure trial, people from the first group were less likely to enter the road early compared to the second group. Kaleefathullah et al. [46] speculate that people of the first group started to rely more on implicit communication of the AV (slowing down), contrary to some participants of the second group who kept relying on the eHMI. The authors acknowledge that their simulation did not feature multiple lanes and, therefore, does not offer results regarding scalability. For a potential failure case in real life, they mention a situation where an AV could detect a pedestrian but would yield to another pedestrian further down the road instead. Our study features conditions similar to this where an AV is yielding for a virtual pedestrian to the participant’s right. We simulated that the AV did not detect the participants crossing intention.

Mahadevan et al. [56] developed a VR traffic simulator where they studied the behavior of pedestrians when confronted with mixed traffic. Some vehicles included users/drivers in either a distracted or attentive state. Additionally, virtual pedestrians were present. Mahadevan et al. [56] found that, generally, pedestrians try to identify the vehicle type (manual, semi-autonomous, fully autonomous) and adjust their crossing behavior. While their data does not show a pattern, they mention that pedestrians could develop an overreliance on predictable AV behavior. Participants stated that the lack of an interface on a vehicle made them more careful. Also, 6 out of 12 participants stated that the behavior of the virtual pedestrians may have influenced their crossing strategy.

In a recent paper, Dey et al. [27] conducted a study where they compared the distance-dependent eHMI solutions proposed by Dey et al. [25] under the aspect of scalability when used in a situation with multiple (2) pedestrians. Participants stood near a road with another virtual pedestrian 10m to their right in their VR setup. An AV was approaching them from the left. It would either drive past the pedestrians or slow down and yield for the participant or the virtual pedestrian. The AV was equipped with one of 4 eHMI solutions (+ one additional *no eHMI* condition): Bumper eHMI, Bumper eHMI + Windshield "Situational Awareness" Display, Bumper eHMI + Windshield "Progress Bar" Display, or Bumper eHMI + Street Projection" Display. Dey et al. [27] measured the participants' willingness to cross the road by instructing them to press and hold a button while the vehicle was approaching until they no longer felt safe to cross. When the AV yielded to the participant, the *Bumper eHMI + Street Projection* performed the best and led to the highest average willingness to cross. The results showed that any of the used eHMI solutions increased the willingness to cross when compared to the *no eHMI* condition. However, the *Bumper + Situational Awareness* concept did not perform better than the *Bumper*. In the scenario where the AV yielded for the virtual pedestrian to the participant's right, the participants showed the lowest willingness to cross when the *Bumper eHMI + Street Projection* was used. The *no eHMI* baseline had a lower willingness to cross than the other eHMI solutions and, therefore, outperformed them in terms of safety. In this specific scenario, the other three eHMI solutions(*Bumper, Bumper + Situational Awareness, Bumper + Progress Bar*) lead to more unsafe behavior. Overall, the *Bumper eHMI + Street Projection* performed best in terms of scalability and was also rated the best experienced by the participants. The authors mention, however, that this eHMI comes with various design challenges as the visual clarity of the projection highly depends on the environment (lighting, weather, road condition, and texture). Their work shows that distance-dependent eHMI concepts do not necessarily perform better in terms of scalability.

Wilbrink et al. [72] conducted a 3x3 within-subject study where videos from the perspective of a pedestrian wanting to cross a one-lane road in a virtual urban environment were shown to participants. An AV was approaching from the left. It was equipped with either no eHMI or a 360° LED-Band around the windshield's edges and extending over the roof to the back of the car that would either communicate the vehicle's automation status or pulsate to indicate the AV's yielding intention in addition to its automation status. Additionally, there was either one virtual pedestrian positioned on the same side as the participant or on the opposite side of the road or no additional pedestrian on the scene. The virtual pedestrian was always closer to the approaching vehicle. The video would play up until a certain point where the participants were asked about their willingness to cross dichotomously (yes/no) and the certainty in their crossing decision using 6-point Likert scales (1 = "very uncertain" to 6 = "very certain"). They found that participants were significantly more willing to cross if the AV communicated its yielding intention. Overall, the willingness to cross was lower if other pedestrians were in the scene. Colley et al. [10] measured a similar effect: Participants rated the 'clarity' of scenarios significantly lower if other pedestrians were present. Beyond that, Wilbrink et al. [72] found no significant effect of the positioning of the additional pedestrian regarding the intention-based eHMI. They claim that the positioning of other pedestrians might not influence its effectiveness. It is important to note that the additional pedestrian was always positioned closer to the approaching vehicle than the participant. The study by Wilbrink et al. [72] appears not to feature any variation in the approaching AV's behavior. It does not state whether or not the vehicle always behaved in the same way (i.e., always yield for the closest pedestrian) or if there was variation (i.e., no yielding, yielding for the first pedestrian, yielding for the second pedestrian). An AV has to predict a pedestrian's crossing intention; therefore, variation in the AV's yielding behavior could be a significant factor in an eHMI's performance regarding its scalability.

Regarding scalability evaluations, mostly VR was used. However, Holländer et al. [44] developed a browser game, 'The Walking Data', to study whether in-game behavior resembles real-world data. Zhuang and Wu [74] found that pedestrians prefer safe paths over short paths when it comes to crossing a road. This matches the results presented by Holländer et al. [44]. Therefore, they argue that games could be used for gathering large-scale data on pedestrian behavior, thus, making it easier and more affordable. Pedestrians mainly tended to change their path in 'safe' areas (sidewalk, exactly in the middle) and crossed the lanes in a straight line. Holländer et al. [44] recognize that cultural differences, and the game's appearance, might impact the results. They suggest that a more photo-realistic approach could influence the behavior.

Dey et al. [22] reviewed numerous eHMI concepts and derived twelve design patterns. These were the basis of our implementation. However, we omitted eyes and other anthropomorphic methods, audio, infrastructure, and mobile devices due to their lower real-world applicability.

3 EXPERIMENT

To evaluate the scalability effects of eHMIs, we designed and conducted a within-subject study with *N*=46. We focused both on the objective behavior logged by the application and the subjective assessment of the eHMIs. This study was guided by the exploratory research questions (RQ):

RQ1 *What impact do the independent variables "eHMI" and "number of pedestrians" have on pedestrians in terms of (1) behavior, (2) mental workload, (3) trust, (4) perceived safety, and (5) acceptance in a complex environment?*

RQ2 *What impact do the independent variables "eHMI" and "automated vehicle behavior" have on pedestrians in terms of (1) behavior, (2) mental workload, (3) trust, (4) perceived safety, and (5) acceptance in a complex environment?*

The study featured two independent variables: The eHMI condition (9) and the pedestrian condition (4). The combination of these leads to a total of 36 different conditions.

We developed a WebGL browser application called "Cross the Road" using Unity (v.2020.3.15f2) [68]. We equipped the AVs with a turquoise LED attached in the center at the top of the windshield as a status indicator as suggested by Faas et al. [32]. Turquoise was used as it is highly visible and has no traffic-relevant meaning [71].

3.1 Environment

The scene for our study is set in a city with two straight two-lane roads. Vehicles enter the area through a tunnel on one side and exit through another on the other side. Participants begin the trial 5m away from the first road. The ground texture changes 2m before the street to a gray texture (indicating a sidewalk, see Figure 1).

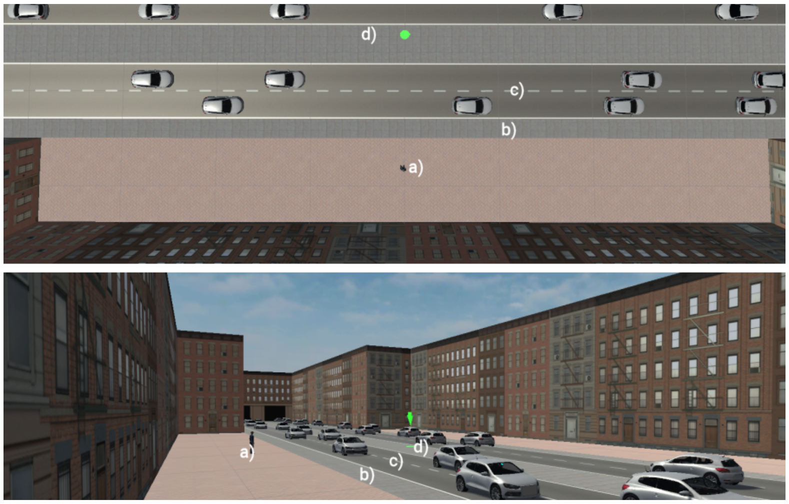


Fig. 1. Setup of each trial. Participants start 5m from the road (*a*). Once they enter the 2m wide sidewalk (*b*), vehicles driving on the road (*c*) can detect and yield for them. Participants had to cross the first road and reach the green waypoint (*d*).

3.2 Pedestrian Conditions

To cover the scalability aspect of multiple pedestrians, we implemented 4 pedestrian conditions. These allow different AV yielding behaviors. Every eHMI is shown to pedestrians 4 times, once per pedestrian condition. The pedestrian conditions do not influence the participant's starting position.

In the Solo conditions, the participant crosses the road alone with no additional virtual pedestrians. This condition was required to enable the evaluation of the effect of pedestrian presence (see Section 4.3).

For all conditions involving multiple pedestrians, two additional virtual pedestrians are present in the trial. They both stand on the sidewalk from the start of the trial 10m apart from each other: One 5m to the left, the other 5m to the right of the participant. We always included both virtual pedestrians to ensure that the participant cannot predict who the AVs are going to yield for based on which virtual pedestrian is present. Depending on the condition, the AVs in both lanes will yield for one of the 3 pedestrians (either the left one, the participant, or the right pedestrian). We 'force' the AVs to misinterpret the behavior of certain pedestrians which prevents the AVs from yielding for them. These cases are similar to what Kaleefathullah et al. [46] describe in their work as 'failure cases'. There is no variation in positioning or rotation of the virtual pedestrians that could give the participant clues about who the AV will yield for.

3.3 Evaluated Concepts

We evaluated nine eHMIs based on eight of the design patterns by Dey et al. [22] (see Figure 2).

3.3.1 Baseline. This condition features modifications to the bumper region of the vehicle. Additionally, a round cyan LED is positioned at the top center of the vehicle's windshield to indicate that the vehicle is driving autonomously. This is meant to help pedestrians differentiate between AVs and manually driven cars [32, 56]. To further indicate that one vehicle is driving autonomously, we added the model of a human sitting inside manually-controlled vehicles while the AV contained no passengers (see Figure 2a). In addition to being one of the nine eHMI conditions, this condition also serves as the *base* for the other eHMI conditions, meaning that they all feature the same modifications. The bumper modification at the bottom represents an LED utilized by the *Bumper Light*, *Situational Awareness*, and *Progress Bar*. Above that, the modification to the grill region of the bumper represents an LED display and is used by the *Bumper Text*, *Pedestrian Symbol*, *Smiling Car*, and *Multi-Lane Bumper Text*. This choice was made deliberately to ensure that differences measured between the eHMI concepts solely rely on the eHMI concept itself. Hence, it serves as a *base* condition.

3.3.2 Bumper Text. This eHMI utilizes the display added to the grill region of the bumper to communicate the vehicle's state via text. While driving, the display says "DRIVING" in bold cyan letters. If the vehicle discovers a pedestrian who wants to cross the road, the text displayed will change to "YIELDING" to communicate its yielding intention until the vehicle stops. The text displayed will then switch to "STOPPED" (see Figure 2b).

3.3.3 Bumper Light. Based on the design by Dey et al. [25], Dey et al. [23], and Dey et al. [27], this eHMI concept features an LED bar added to the bumper, which is used to communicate the AV's state and intention. While the AV is driving, the LED bar is glowing statically in cyan. When yielding for a pedestrian, the bar starts to pulsate between on and off at a rate of 1 Hz. The pulsation is based on work by Dey et al. [23] who showed that a cyan "a uniformly flashing or pulsing animation is preferred compared to any pattern that animates sideways" [23, p. 1]. While Mirnig et al. [58] were able to show that a green-red traffic light metaphor improved interaction success, the red-green metaphor was omitted to avoid giving the appearance of being a traffic light. Once the AV comes to a full stop, the light bar turns off (see Figure 2d). This 1-dimensional light bar design [22] does not change its behavior based on the *pedestrian conditions*. Like in previous works [25, 27], this eHMI is used in combination with the distance-dependent eHMI concepts *situational awareness* and *progress bar*.

3.3.4 Situational Awareness Indicator. In addition to using the *bumper light*, this concept - falling into the "Tracking Light" design pattern [22] category - is based on the design of Dey et al. [25, 27]: a cyan indicator on the windshield of the vehicle points toward the pedestrian the AV intends to yield for (see Figure 2c). Contrary to the concept by Dey et al. [25], in which the indicator is a rectangle ranging from the bottom to the top of the

(a) *base* (left, with cyan LED). Right: manual control.

(b) The *Bumper Text* communicates the current AV state.

(c) *Situational awareness. a)* shows the AV yielding. The indicator follows the pedestrian (*b*), *c)*).

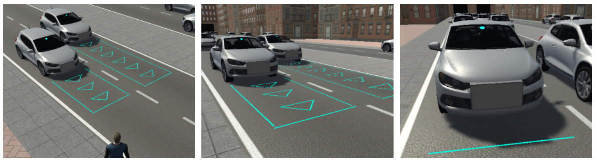
(d) *Bumper Light*: while driving (left), pulsating at 1 Hz when yielding (middle), and off when AV stops (right).

(e) Distance-dependent *progress bar* [25, 27].

(f) *Pedestrian symbol.*

(g) *Smiling car* [5, 64].

(h) The *multi-lane bumper text.*



(i) The *street projection a.* It moves with the AV until 10m of its yielding position. Then, the end of the projection shows the stopping position (*b*). The AV drives over its projection and "erases" it (*c*). Finally, only the line remains (*d*).

Fig. 2. The employed feedback strategies used in the experiment.

windshield, we utilized a square at the bottom of the windshield as this implementation is more realistic. The concept by Dey et al. [25] requires a full windshield display [53] whereas our solution could be implemented using an LED band at the bottom of the windshield. We also decided to add additional behavior when the vehicle detects multiple pedestrians intending to cross. In the study conducted by Dey et al. [27], one additional virtual pedestrian was added besides the participant. The AV's indicator would point towards the pedestrian it is going to yield for. Because our study features two extra pedestrians in some of the pedestrian conditions, our eHMI would display one indicator for each pedestrian it detected that has the intention to cross the road. Dey et al. [27] measured the participants' willingness to cross without giving them the option to actually step into the road. Consequently, we further adjusted the eHMI's behavior: If the participant steps onto the road even though the vehicle did not recognize their intention to cross, a new indicator would now point towards the participant. If the participant is closer to the vehicle than its intended stopping position, the previously existing indicator will point towards them instead of the original yielding target. Additionally, after an AV stopped for a pedestrian, it will add an indicator for every other pedestrian that walks onto the road.

3.3.5 Progress Bar.

The progress bar is based on the works of Dey et al. [25, 27]. In combination with the *Bumper Light*, it features an abstract light-based display design pattern [22] by displaying a progress bar that fills the entire windshield. When the yielding process starts, it appears as a small cyan vertical bar in the middle on the windscreen that expands horizontally to the left and the right based on the distance to the stopping point: The shorter the distance, the wider the progress bar (see Figure 2e). Once it stops, the progress bar covers the AV's entire windshield. It remains visible once the vehicle stopped (see Figure 2e). Because participants were not given the freedom of movement in the studies by Dey et al. [25, 27], the concept had to be updated to work with our study design: If the AV was required to update its yielding position because the participant walked towards it, the progress bar would update and fill up based on its new yielding position. The progress is still calculated based on the AV's position at the beginning of the yielding process rather than its current location to avoid the progress bar from fully resetting back to zero and filling up rapidly.

3.3.6 Pedestrian Symbol.

This eHMI uses a symbol design pattern [22] and is based on the concept by Clamann et al. [7]. In their study, the vehicle would communicate if it is safe to cross or not via a pedestrian street symbol on a display attached to the vehicle's grill. By default, the pedestrian displayed was crossed out, indicating that it is not safe to cross. If a vehicle intended to yield for a pedestrian, the symbol would no longer be crossed out to communicate to the pedestrian that it is now safe to cross. The symbol was chosen deliberately to mimic an American street sign. Our implementation varies slightly: By default, the display doesn't show any additional information and only shows it once it yields by showing the symbol of a pedestrian (see Figure 2f).

3.3.7 Smiling Car.

Our implementation of the Smiling Car by Semcon [5] is based on how the video depiction [64]. While being in its own design pattern category in the eHMI taxonomy by Dey et al. [22], this concept can be described as anthropomorphic. It is the only anthropomorphic eHMI we implemented because, contrary to other concepts [56, 57, 63], it can be implemented without drastically changing the design of the vehicle. The eHMI uses the AV's bumper display and draws a mouth on it. If the AV is not yielding, the mouth is represented as a horizontal line: a neutral expression. If the AV is yielding for a pedestrian, the mouth turns from the horizontal line into a smile (see Figure 2g). No additional modifications were done to address multiple pedestrians. The idea is that the AV tells the participant that it has noticed them and that they can cross.

3.3.8 Multilane Bumper Text.

We implemented an additional version of the *Bumper Text* to experiment with an eHMI where only one vehicle acts as a representative of all AVs that intend to yield for the pedestrian. Colley et al. [15] describe this concept as an 'omniscient narrator', which was deemed especially useful for people with vision impairments. The more lanes there are, the more AVs would display information. Additionally, The more lanes a road has, the further away the vehicle is, and, thus, the harder it is for a pedestrian to read and interpret the eHMI due to a steeper angle. Finally, too many eHMIs communicating individually means more information needs to be processed by the pedestrian, which could lead to information overload. Therefore, in the 'omniscient narrator' concept [15], by default, each AV communicates its state by itself. While driving normally, the bumper display will say "DRIVING". As soon as an AV begins to yield, the text will switch to "YIELDING". If vehicles are yielding (or have already stopped) on both lanes, the display of the AV on the lane closest to the pedestrian will switch to "ALL YIELDING" to indicate that the vehicles are yielding on all lanes. The AVs on the other lane will turn off their text. To ensure that no misinterpretations can occur, this will only happen if the distance between the AVs (each of them on their lane) is less than 12m. Otherwise, there could be a scenario where one AV has already stopped on the lane closest to the waiting pedestrian, with another AV beginning its yielding process on the second lane. In this case, there is a chance that another manually controlled vehicle does not intend to yield in front of the AV on the second lane that has not passed the pedestrian yet. The AV in the first lane could obstruct the vision of the pedestrian. If it displayed "ALL YIELDING", this could lead to the pedestrian stepping in the second lane thinking it is safe, leading to an emergency break in the best and a collision in the worst case. This concept requires cross-compatibility between AVs of different manufacturers.

3.3.9 Street Projection. This eHMI concept performed best among the distance-dependent eHMIs in the study by Dey et al. [27] which was based on the Mercedes F015 [4] and the works of Löcken et al. [52] and Dietrich et al. [29]. However, we do agree with Dey et al. [27] that the eHMI could be deemed unrealistic as it requires near-perfect environmental conditions regarding lighting and road conditions to work properly. Also, a powerful projector would be required because the eHMI is projected over a large distance. Therefore, we decided to alter the design to be considered more realistic: Once the AV begins to yield, it will project a rectangle with triangles representing arrows up to 10m in front of the car. The projection will move with the vehicle until it is 10m away from its yielding position. At this point, the end of the rectangle represents the line that marks the spot where the vehicle will stop. The vehicle will continue to come closer to its stopping position and thus decrease the distance between itself and the yielding line, which creates an "erasing" effect when the AV is driving over its projection (see Figure 2i). Once stopped, the stopping line will remain present. We decided against turning the projection into a zebra crossing [4, 27] as a zebra crossing gives pedestrians certain rights and is an official street sign. The eHMI would adjust its projection if the AV is forced to yield at a position closer to the intended yielding position. Also, contrary to Dey et al. [27], our version of the street projection doesn't include a *bumper light*.

3.4 Measurements

3.4.1 Objective Measurements. The objective dependent variables were logged by the application. The application logged the participant's position at 1Hz. Additionally, the time and position of entering a lane (both near and far) and the total time to cross were logged. Also, the application logged the yielding position and the vehicle's stopping position. This log included the vehicle type (AV vs. Non-AV), the vehicle's lane, and the ID of the pedestrian it yielded for. We also logged the yielding reason, which was divided into:

0: Voluntary If an AV detects a pedestrian on the sidewalk, it will yield for the pedestrian.
1: Yielding Adjustment If an AV is already yielding for a participant but is forced to adjust its yielding position by the participant walking towards it
2: Involuntary Yield The default case for Non-AVs as they never intend to yield for a pedestrian. This is also the case if an AV intends to yield for someone else but the participant steps in the vehicle's path.
3: Emergency Break If a participant steps in front of the vehicle 4m in front of it, the vehicle will attempt a so-called emergency break. 4m was chosen as this allows an emergency break from approximately 30 km/h to a standstill and can generally be considered as dangerous behavior.

Based on the lane of the yielding log, we take the LaneEnter log closest to the start of the yielding process: $LaneEnterDifference = TimeOfVehicleStopping - TimeOfPlayerEnteringTheLane$ If the $LaneEnterDifference$ is negative, the participant entered the road before the vehicle stopped. Finally, collisions were also logged, including the vehicle, the user ID, the position, and the time.

3.4.2 *Subjective Measurements.*

The subjective dependent variables were collected via an in-game questionnaire. After every trial, we employed the raw mental workload subscale of the raw NASA-TLX [40] 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* (*Understanding* from here) and *Trust* of the *Trust in Automation* questionnaire by Körber [47]. Understanding 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."). Also, participants rated their perceived safety using four 7-point semantic differentials from -3 (anxious/agitated/unsafe/timid) to +3 (relaxed/calm/safe/confident) [31]. Additionally, we employed the van der Laan acceptance scale [69] with the subscales "usefulness" and "satisfying".

After encountering all 36 trials, participants were asked to rank the nine eHMI conditions regarding their preference. They were also asked about the *Necessity* and *Reasonability* of eHMI concepts using a 7-point Likert scale ranging from *Totally Disagree* to *Totally Agree*. Finally, participants could provide open feedback.

3.5 Procedure

The study started with a short survey on LimeSurvey to introduce the study requirements and context. After asking for their consent to the processing of their data in conformity with the General Data Protection Regulation (GDPR), participants were asked to enter a user name. Participants then received a link to 'Cross the Road'.

First, the game's task was explained in text form: They take the role of a pedestrian who wants to jaywalk across a busy road with mixed traffic. After consenting to the processing of their data in conformity to the GDPR a second time, entering their user name from before, and answering demographic questions, participants were put into a small tutorial area where the movement mechanics - walking using the WASD keys or the Arrow keys and looking around using the mouse - were explained to them. They were instructed to walk towards a green arrow symbol once they felt comfortable with the controls. This brought them to the actual study. Upon the start of a trial, the player was met with the prompt: "Try to cross the road safely and reach the green waypoint". The order the participants encountered the conditions was automatically assigned by the application and is based on a balanced Latin Square. After every crossing, participants encountered the in-game questionnaires. After all 36 trials, a code was shown to the participants, which they were instructed to use in the LimeSurvey survey. After entering the code, they were asked questions about the necessity of eHMIs and ranked the 9 eHMI conditions. The study took, on average, 45 min. Participants were compensated with 4.73€.

4 RESULTS

4.1 Data Analysis

We evaluated the effect of *the number of pedestrians* and the *behavior of automated vehicles* separately, as applying the yielding behavior, for example, yielding after the participant would have to be considered an erroneous behavior of the AV when only the participant is present.

Prior to statistical tests, we checked the required assumptions (normality distribution and homogeneity of variance assumption). For non-parametric data, we used the non-parametric ANOVA (NPAV) [55]. For post-hoc tests, we used Dunn's test with Bonferroni correction. We employed R in version 4.2.2 and RStudio in version 2022.12.0. All packages were up to date in November 2022.

4.2 Participants

We determined the required sample size via an a-priori power analysis using G*Power in version 3.1.9.7 [34]. To achieve a power of .95 with an alpha level of .05, 38 participants should result in an anticipated small to medium effect size (0.15 [37]) in a within-factors repeated-measures ANOVA with two groups and 18 measurements.

$N=46$ participants (22 male, 24 female, 0 non-binary, 0 undecided) with an average age of $M=29.56$ ($SD=10.20$) years took part in the study. As their highest educational level, 28 out of 46 participants listed "college", 16 "High School", 1 "Vocational Training", and 1 "Graduate school". 17 participants described themselves as "employee", 13 as "college student", six as "self-employed", eight as "job-seeking", and two as "unemployed". When asked about their interest in autonomous driving using a 5-point Likert scale (1 = *Not at all*, 5 = *Definitely*), participants showed a slight interest ($M=3.56$, $SD=1.09$). Using the same Likert scale, participants, on average, stated that AVs will make their life easier ($M=3.83$, $SD=1.08$). They also believe, on average, that AVs will become a reality within the next 10 years ($M=4.13$, $SD=0.96$).

4.3 Effect of Pedestrian Presence

This section presents the results to RQ1 *What impact do the independent variables “eHMI” and “number of pedestrians” have on pedestrians in terms of (1) behavior, (2) mental workload, (3) trust, (4) perceived safety, and (5) acceptance in a complex environment?*.

4.3.1 Mental Workload.

The NPAV found a significant main effect of *eHMI* on mental workload ($F(8,360) = 2.53$, $p=0.011$). However, a post-hoc test did not show any significant differences.

4.3.2 Trust in Automation.

Understanding. For *Understanding*, a NPAV found a significant main effect of *eHMI* ($F(8,360) = 3.47$, $p<0.001$). A post-hoc test revealed that *Understanding* for the *Baseline* condition ($M=3.65$, $SD=1.22$) was significantly lower than the *Pedestrian Symbol* ($M=4.30$, $SD=0.87$, $p=0.007$) and the *Street Projection* ($M=4.18$, $SD=1.01$, $p=0.045$). The NPAV found no significant effect of *pedestrians* on *Understanding*.

Trust. The NPAV found a significant main effect of *eHMI* on *Trust* ($F(8,360) = 2.36$, $p=0.017$). A post-hoc test, however, did not show any significant differences. The NPAV found no significant effect of *pedestrians* on *Trust*.

4.3.3 Acceptance.

Usefulness. The NPAV found a significant main effect of *eHMI* on usefulness ($F(8,360) = 3.59$, $p<0.001$). The post-hoc test showed that the *Usefulness* of the *Baseline* ($M=0.53$, $SD=1.18$) was rated significantly worse than the *Pedestrian Symbol* ($M=1.27$, $SD=0.75$, $p<0.001$), the *Progress Bar* ($M=1.05$, $SD=0.95$, $p=0.037$) and the *Street Projection* ($M=1.22$, $SD=0.84$, $p<0.001$).

The NPAV found a significant main effect of *pedestrians* on usefulness ($F(1,45) = 4.27$, $p=0.045$). However, a post-hoc test did not show any significant differences.

Satisfying. The NPAV found a significant main effect of *eHMI* on *Satisfying* ($F(8,360) = 2.91$, $p=0.004$). A post-hoc test revealed that *Satisfying* was rated significantly lower for the *Baseline* condition ($M=0.48$, $SD=1.24$) when compared to the *Pedestrian Symbol* ($M=1.13$, $SD=1.06$, $p=0.009$) and the *Street Projection* ($M=1.18$, $SD=0.97$, $p=0.004$) eHMI condition. The NPAV found no significant effect of *pedestrians* on *Satisfying*.

4.3.4 Perceived Safety.

The NPAV found a significant main effect of *eHMI* on perceived safety ($F(8,360) = 2.15$, $p=0.030$). a post-hoc test did, however, not show any significant differences. The NPAV found no significant effect of *pedestrians* on perceived safety.

4.3.5 Road Crossing Behavior. We only included where AVs yielded in both lanes for this analysis. No dangerous collisions occurred.

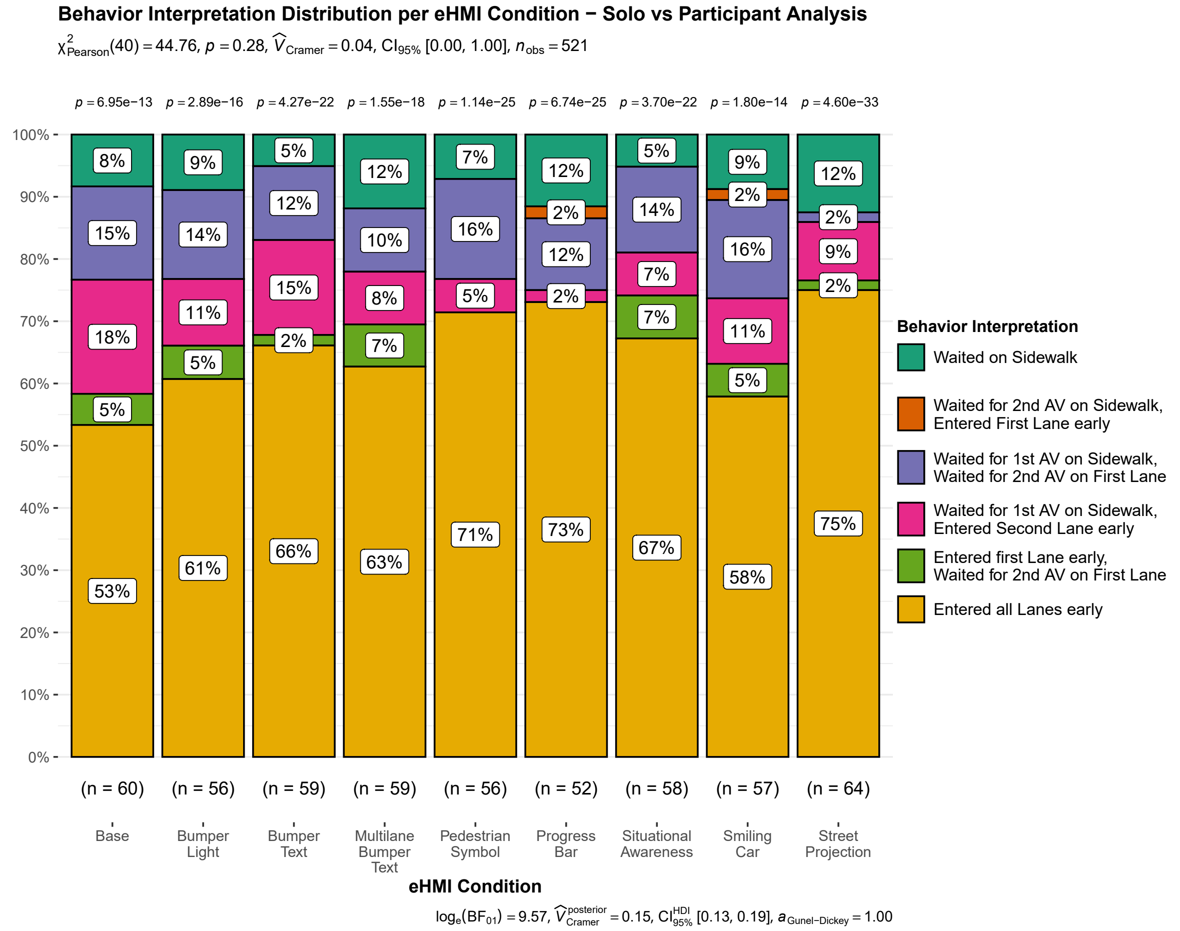


Fig. 3. Behavior Interpretation Distribution per eHMI Condition - Mere Presence Effect Analysis. For all eHMI conditions, in most cases, participants entered the lanes before any vehicle entirely stopped, with the highest rate being 75% of the *Street Projection* eHMI. This eHMI also has the highest rate of participants waiting for both vehicles to fully stop at the sidewalk (12%) and the lowest rate (2%) of participants waiting for the first AV on the sidewalk and for the second AV on the first lane.

We analyzed the road crossing behavior of 521 game logs and found that in most cases (340; 65.26%), participants entered both lanes early, meaning they did not wait for any of the two AVs to completely stop before crossing. In 63 out of 521 cases (12.1%), participants waited on the sidewalk for the first vehicle to stop and then proceeded to wait for the second vehicle to stop in the first lane. In 51 cases (9.8%), participants waited on the sidewalk for the first vehicle to stop but didn't wait for the second AV to stop. In 46 cases (8.8%), participants waited on the sidewalk for both vehicles to stop. In 19 cases (3.6%), participants entered the first lane early before the AV stopped but then waited for the AV in the second lane to stop fully. In 2 cases (0.4%), the AV in the second lane arrived before the AV in the first lane, and participants entered the first lane before the AV in the first lane fully stopped. Looking at the distribution of the behavior interpretations across the different eHMI conditions, crossing before any vehicle stopped makes up the majority for every eHMI condition (see Figure 3) with the *Baseline* condition having the lowest rate of early lane enters (53%) and the *Street Projection* having the highest with 75%. At the same time, the *Street Projection* eHMI has the highest rate (12%) of people waiting on the sidewalk for both vehicles to stop together with the *Multilane Bumper Text* and *Progress Bar*. When comparing the pedestrian conditions, there was no significant difference between the *Solo* and the *Pedestrian* condition.

4.4 Effect of Yielding Behavior

This section presents the results to RQ2 *What impact do the independent variables "eHMI" and "automated vehicle behavior" have on pedestrians in terms of (1) behavior, (2) mental workload, (3) trust, (4) perceived safety, and (5) acceptance in a complex environment?*.

4.4.1 Mental Workload. The NPAV found a significant main effect of *eHMI* on mental workload ($F(8, 360) = 2.57$, p=0.010). The post-hoc test showed that *mental workload* for the *Baseline* condition (M=6.63, SD=6.19) was significantly higher when compared to the *Pedestrian Symbol* (M=3.62, SD=4.80, p=0.002) and the *Progress Bar* (M=4.25, SD=5.22, p=0.049). The NPAV also found a significant main effect of *pedestrians* on mental workload ($F(2, 90) = 15.04$, p<0.001). The *mental workload* for the *Right* pedestrian condition (M=5.64, SD=6.01) was significantly higher than the *Left* condition (M=4.23, SD=5.25, p=0.003).

4.4.2 Trust in Automation.

Understanding. The NPAV found a significant main effect of *eHMI* on *Understanding* ($F(8, 360) = 4.73$, p<0.001). The post-hoc test showed the *Baseline* condition (M=3.58, SD=1.22) was rated significantly worse than the *Bumper Text* (p=0.003), the *Pedestrian Symbol* (M=4.24, SD=0.88, p<0.001), the *Progress Bar* (M=4.02, SD=1.14, p=0.032) and the *Street Projection* condition (M=4.05, SD=1.14, p=0.008).

The NPAV found a significant main effect of *pedestrians* on understanding ($F(2, 90) = 15.54$, p<0.001). The *Right* pedestrian condition (M=3.75, SD=1.22) was rated significantly worse than the *Participant* pedestrian condition (M=4.06, SD=1.04, p=0.003) and the *Left* pedestrian condition (M=4.18, SD=0.97, p<0.001).

Trust. The NPAV found a significant main effect of *eHMI* on trust ($F(8, 360) = 2.63$, p=0.008). Trust was rated significantly lower for the *Baseline* (M=3.70, SD=1.25) in comparison to the *Pedestrian Symbol* (M=4.22, SD=0.96, p=0.017) and the *Street Projection* (M=4.09, SD=1.23, p=0.037).

Additionally, the NPAV found a significant main effect of *pedestrians* on trust ($F(2, 90) = 17.61$, p<0.001). Trust was rated significantly lower for the *Right* pedestrian condition (M=3.80, SD=1.25) in comparison tot he *Left* (M=4.19, SD=1.05, p<0.001) and the *Participant* (M=4.07, SD=1.09, p=0.012) pedestrian condition.

4.4.3 Acceptance.

Usefulness. The NPAV found a significant main effect of *eHMI* on usefulness ($F(8, 360) = 3.95$, p<0.001). The *Baseline* (M=0.47, SD=1.19) was rated significantly less useful than the *Bumper Light* (M=1.01, SD=0.92, p=0.007), *Bumper Text* (M=1.09, SD=0.81, p=0.002), *Multilane Bumper Text* (M=0.97, SD=0.94, p=0.021), *Pedestrian Symbol* (M=1.25, SD=0.78, p<0.001), *Progress Bar* (M=1.01, SD=0.95, p=0.002), *Situational Awareness* (M=0.99, SD=0.94, p=0.007), *Smiling Car* (M=0.82, SD=1.20, p=0.043) and *Street Projection* eHMI (M=1.13, SD=0.88, p<0.001).

The NPAV found a significant main effect of *pedestrians* on usefulness ($F(2, 90) = 14.94$, p<0.001). The *Left* (M=1.10, SD=0.91) was rated significantly better than the *Right* pedestrian condition (M=0.84, SD=1.05, p=0.001).

Satisfying. The NPAV found a significant main effect of *eHMI* on satisfying ($F(8, 360) = 3.67$, p<0.001). A post-hoc test showed that the *Base* (M=0.36, SD=1.28) was rated significantly worse than the *Bumper Light* (M=0.85, SD=1.16, p=0.040), the *Pedestrian Symbol* (M=1.12, SD=1.00, p<0.001), the *Progress Bar* (M=0.87, SD=1.24, p=0.014), the *Smiling Car* (M=0.92, SD=1.20,p=0.003) and the *Street Projection* (M=1.06, SD=1.05,p<0.001). The NPAV found a significant main effect of *pedestrians* on satisfying ($F(2, 90) = 9.94$, p<0.001). The *Left* pedestrian condition (M=0.97, SD=1.11) was rated significantly better than the *Right* pedestrian condition (M=0.70, SD=1.23, p=0.004).

4.4.4 Perceived Safety. The NPAV found a significant main effect of *eHMI* on perceived safety ($F(8, 360) = 3.08$, p=0.002). The post-hoc test shows that the *Perceived Safety* of the *Baseline* condition (M=1.22, SD=1.68) was rated significantly lower than the *Bumper Text* condition (M=1.85, SD=1.48, p=0.027).

The NPAV found a significant main effect of *pedestrians* on perceived safety ($F(2, 90) = 12.62$, p<0.001). The post-hoc test shows that *Left* pedestrian condition (M=1.79, SD=1.48) was rated significantly better than the *Right* pedestrian condition (M=1.24, SD=1.80, p<0.001).

4.4.5 Road Crossing Behavior. For this analysis, we only included where AVs yielded in both lanes for this analysis, and no dangerous collisions occurred.

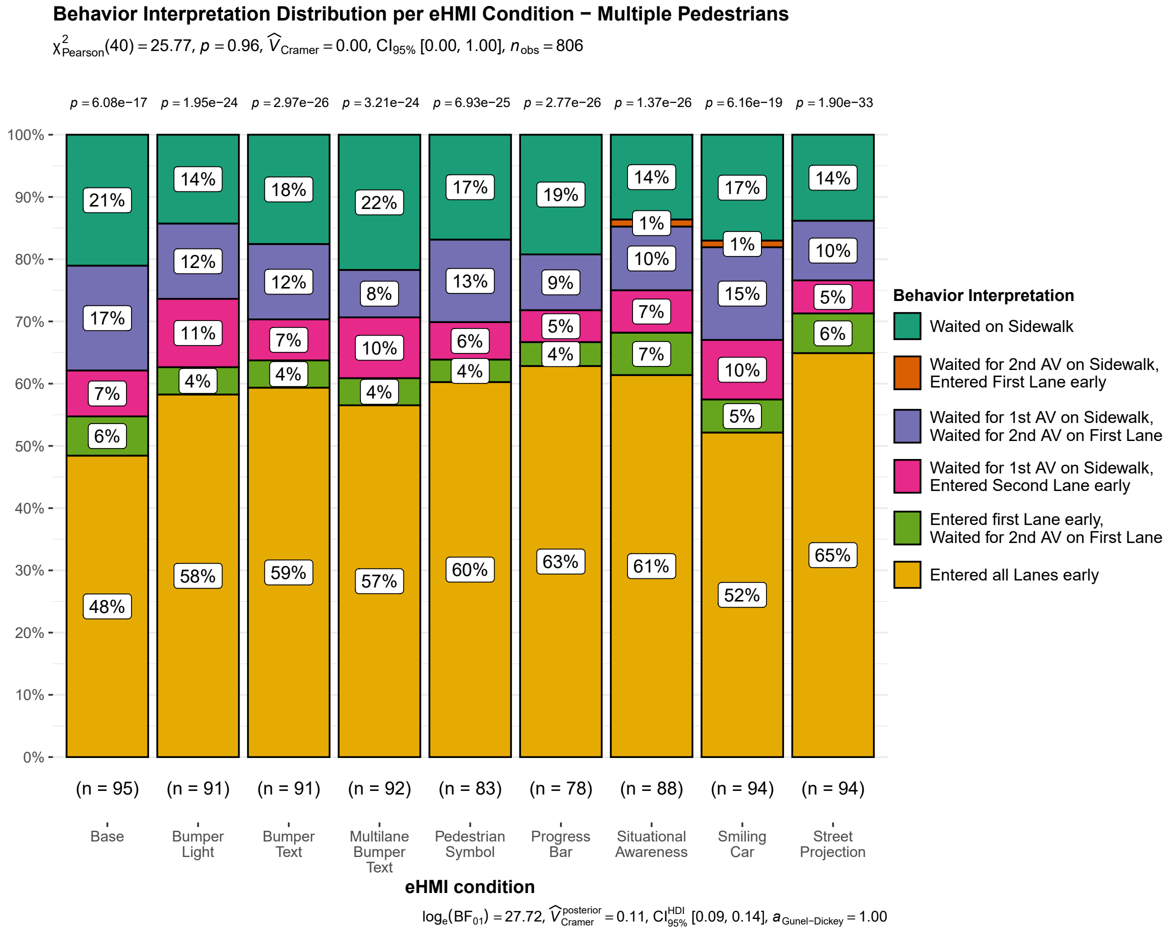


Fig. 4. Behavior Interpretation Distribution per eHMI Condition - Effect of Yielding Target Analysis

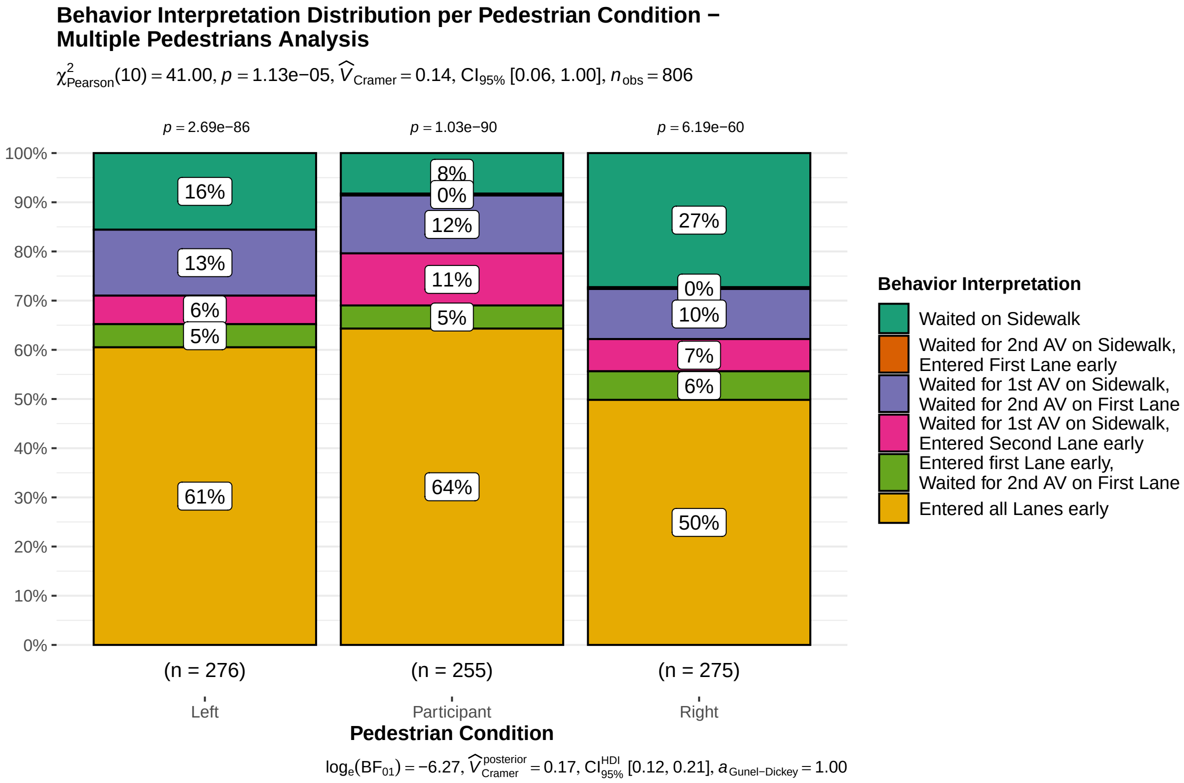


Fig. 5. Behavior Interpretation Distribution per Pedestrian Condition - Effect of Yielding Target Analysis

We analyzed the road crossing behavior of 806 game logs. We found that in most cases (468, 58.06%), participants entered botWeg they did not wait for any of the two vehicles to stop completely before crossing. In 139 cases, participants waited for both vehicles to stop on the sidewalk. In 95 cases, participants waited on the sidewalk for the first AV to stop and then waited in the first lane for the second AV to stop. In 61 cases, participants waited on the sidewalk for the first AV to stop but crossed the road before the second vehicle stopped. In 41 cases, participants entered the road before the first AV fully stopped but then waited in the first lane for the AV in the second lane to stop. In 2 cases, the AV in the second lane arrived before the AV in the first lane, and participants entered the first lane before the AV in the first lane fully stopped. Looking at the distribution of the behavior interpretations across the different eHMI conditions, crossing before any vehicle stopped makes up the majority for every eHMI condition (see Figure 4) with the *Baseline* condition having the lowest rate of early lane enters (48%). The *Street Projection* having the highest with 65%. The *Multilane Bumper Text* had the highest rate of participants waiting at the sidewalk (22%), followed by the *Baseline* (21%) and the *Progress Bar* (19%). When comparing the pedestrian conditions, participants waited on the sidewalk for 27% of the cases involving the *Right* pedestrian condition (see Figure 5).

4.5 Collisions

A total of 305 collisions occurred across all game logs, 143 *Dangerous* collisions and 162 *Safe* collisions (the participant touched a vehicle that was not moving). 220 collisions involved Non-AVs, 85 AVs. 98% of the *Dangerous* collisions involved Non-AVs (49% for *safe* collisions). Looking at the eHMI condition distribution, most *dangerous* collisions (15%) occurred during trials involving the *situational awareness* and the *Progress Bar*. Looking at the pedestrian conditions, 50% of *safe* collisions were recorded in trials with the *right* pedestrian condition.

By filtering the gameplay logs using the same parameters used for the *Yieldings* (AVs yielded in both lanes) but including dangerous collisions, we found 86 collisions, of which only 3 were dangerous collisions, two of which involved AVs. Figure 7 shows a visualization of the collision location. Looking at the *Safe* collisions (participants touching a vehicle that is not moving), most collisions occurred with AVs. Additionally, the majority of *safe* collisions occurred during *Right* pedestrian conditions trials. The majority of *Safe* collisions during the *Right* pedestrian occurred on the right side where participants had to walk around the AV (see Figure 8).

4.6 Ranking

Participants were asked to rank the nine eHMIs based on which they liked the best (The lower the rank, the better; see Figure 6). With an average rank of *M*=3.13, the *Pedestrian Symbol* was rated best, followed by the *Street Projection* (*M*=3.37) and the *Bumper Light* (*M*=3.70). A Durbin-Conover test found no statistically significant difference between those three eHMI concepts. The *Baseline* condition was rated the worst with an average rank of *M*=8.24, followed by the *Situational Awareness* (*M*=7.17).

4.7 Necessity, Reasonability, and Open Feedback

On average, participants stated that there is a necessity for eHMIs (*M*=5.98, *SD*=1.64) and that the use of eHMIs is reasonable (*M*= 6.35, *SD*=1.35).

Regarding the *Bumper Light*, the *Pedestrian Symbol*, and the *Smiling Car*, participants mentioned that they liked that it was easy to recognize from a distance. One participant mentioned that they especially liked the *Pedestrian Symbol* because it used a known road sign. Also, contrary to others, one participant deemed the *Smiling Car* as pointless. Multiple participants mentioned that they had difficulties reading the text-based eHMIs and that people with vision impairments could have difficulties reading the text in real life. Additionally, while one participant mentioned that they liked the additional information provided by the *Multilane Bumper Text*, multiple participants deemed the concept confusing or described it as "buggy". While some participants pointed out that the *Progress Bar* was more predictable, one participant mentioned that they did not feel safe to cross because they could not see the vehicles in the second lane. One participant pointed out that eHMI communication needs to be standardized and easy to understand for people of all ages. In general, many participants felt that the study was too long, and the AV's yielding behavior became more predictable with time.

5 DISCUSSION

Overall, all investigated eHMI concepts were preferred compared to no eHMI. In line with previous work, these concepts improved mental workload, trust, and crossing onset time. Despite the overall benefits of eHMIs, there were still clear preferences regarding the concepts. Additionally, we discuss the results in the light of safe crossing behavior and guide future developers in their choice of eHMI.

5.1 Ecological Validity of Study Approach

We employed a WebGL-based browser game to study the effects of eHMIs, the number of pedestrians, and AV behavior. Numerous evaluations of eHMIs and their effects have been studied in VR [9, 12, 48], CAVE simulators [50], or by employing video both of filmed real vehicles [19, 21] or simulated ones [11, 39]. Lew et al.

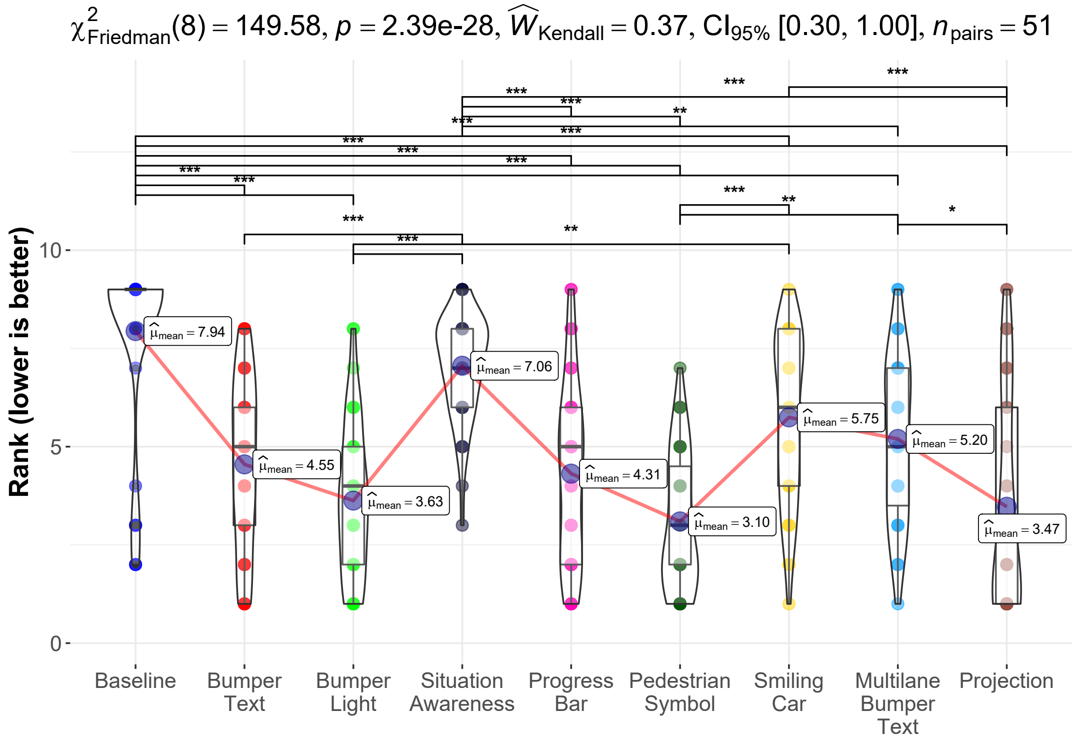


Fig. 6. Results of the eHMI ranking.

[51] discuss the four areas of realism appearance, content, task, and setting to confound the study's validity. By employing a realistic appearance of the vehicles, pedestrians, and other content, a known task (street crossing), and a known setting (urban city), we believe to provide sufficient realism for the study. Previously, browser-based games have also been used to teach children crossing rules [8]. This work showed that the use of this game increased real-world behavior [8], increasing the argument for the transferability of the findings. Dong et al. [30] compared pedestrian gaze in desktop and real environment. They "generally found more differences than similarities between the street view and real world" [30, p. 17]. The authors attribute this, for example, to the increased noise levels in the real world. While eye gaze was no objective in our study, this diminishes the argument for transferability.

Fuest et al. [36] compared videos (of a real vehicle and a simulated one) with VR and a Wizard-of-Oz vehicle. While the authors report some limitations of their study (e.g., the VR scene differed significantly from the real-world scenario and lacked sound), the authors showed significant but descriptively small differences between the conditions. They especially highlight that there might be "perception and decision artefacts" [36, p. 22] in the video conditions, but those small deviations cannot be evaluated in the Wizard-of-Oz approach due to human failures. This work's findings results lower the argument for transferability.

Until now, one work employed a browser-based game to evaluate crossing scenarios [44]. For this, they employed a low-fidelity environment. However, they found that the observed walking behavior is comparable to real-world behavior [44] by comparing their data with data from China provided by Zhuang and Wu [75].

We view this comparison as the most influential argument for the transferability of our findings to the real world. Nonetheless, the results for objective variables of our experiment, such as crossing time, depended on the participant's decision and the chosen implementation parameters (e.g., walking speed). While this improved internal validity, external validity was negatively impacted. Additionally, we chose the study's setup with vehicles of the same color and size to increase internal validity. This again comes with the drawback of reduced ecological validity. Therefore, this study should be evaluated using VR or the real world.

5.2 Interpreting Road Crossing Behavior

Pedestrian crossings should be as safe as possible while spending the least time on the road. The paths participants took varied mainly during *Right* pedestrian condition trials because AVs stopped to the right of the participant. Thus, players either had to cross behind the AV or in front of it. The 'curve' that can be observed suggests that pedestrians prefer safe crossing behavior over fast crossing behavior and that the game was performed realistically (in line with [44, 74]). While this indicates that road crossing behavior **can** be studied by using a browser game, future work is necessary to validate this assumption.

Wilbrink et al. [72] concluded that the presence of other pedestrians alone reduces the 'Willingness to Cross'. Colley et al. [10] also measured lower 'clarity' in scenarios with additional pedestrians present. Our *Mere Presence Effect* analysis (see Section 4.3) showed no significant differences in *Road Crossing Behavior* whether or not multiple pedestrians are present. In our study - where vehicles come from the left - the *Right* pedestrian was the only condition where pedestrians were subject to danger when misinterpreting the AVs intention as the vehicle would drive past them to yield for the pedestrian to their right. The *Behavior Interpretation* shows that this condition had the overall lowest rate of early lane enters of any kind while having the highest rate of participants waiting on the sidewalk (low 'Willingness to Cross'/'clarity'). These findings agree with Dey et al. [27] where participants also showed the lowest 'Willingness To Cross' if the vehicle stopped for a virtual pedestrian further down the road. This 'failing case' [46] also had the lowest rate of *Voluntary* yieldings (58%) and the highest rate of changed yielding behavior by the AVs. In addition, participants felt significantly less satisfied and safe compared to the *Left* pedestrian condition. This indicates that future research needs to take the *Effect of Yielding Target* into account rather than the *Mere Presence Effect* alone. When investigating the behavior interpretations of the eHMIs combined with the *Right* pedestrian condition, the *Baseline* condition had by far the lowest rate of early lane enters and the highest rate of pedestrians waiting for vehicles to stop either on the sidewalk or in the first lane after the first AV stopped. These results are similar to the *Behavior Interpretation* including all pedestrian conditions. These findings are somewhat contradictory to Dietrich et al. [29] as the undirected eHMIs (*Base*, *Bumper Light*, *Bumper Text*, *Multilane Bumper Text*, *Pedestrian Symbol*, *Smiling Car*) all had higher rates of 'safe' crossing behavior than the directed eHMIs (*Situational Awareness*, *Progress Bar*, *Street Projection*). However, here, early crossing behavior could also imply that participants understood the situation earlier and decided to cross earlier. This contradicts the study results by Clamann et al. [7] who found that eHMIs had no impact on the 'decision time to cross'.

5.3 Performance of External Human-Machine Interfaces

The results for the *Multilane Bumper Text* eHMI concept are somewhat inconclusive. As many participants misunderstood the concept underlines that, in hindsight, the eHMI maybe should have indicated that another car is 'speaking' for them. At this point, it raises the question of whether an 'omniscient narrator' [15] or 'representative' communicating the intention of all AVs is a good idea in the first place or potentially only for a subset of road users (i.e., people with vision impairments).

Similar to the study by Dey et al. [27], our iteration of the *Street Projection* eHMI was among the best performing concepts. As it was already mentioned in their work, we agree that this concept might be unrealistic as it relies on perfect environmental conditions regarding lighting, weather, time of day, and road condition to work properly. Similar to their work [27], both the *Progress Bar* and *Bumper Light* were also among the best-rated eHMIs and performed both better than the *Situational Awareness*. Interestingly, the *Bumper Light* condition does not appear as a significant factor in any of the analyses, neither for better nor for worse. This might indicate that its use in other eHMI conditions (*Progress Bar*, *Situational Awareness*) might have influenced participants' ranking decisions. Our modification to the *Situational Awareness* eHMI did not help to improve its performance under the aspect of scalability. As it had the second-highest rate of participants waiting at the sidewalk for both AVs to stop under the *Right* pedestrian at 31% compared to having the lowest rate across all pedestrian conditions (12%), it shows that pedestrians failed to identify whether or not the AV was communicating with them or the other pedestrian. The same applies to the *Bumper Text* eHMI.

The *Smiling Car* was rated poorly, showing that the concept failed to perform well in a scaled mixed-traffic scenario. As it intended to communicate whether or not it is safe to cross like the *pedestrian symbol* (rated best), these results imply that the concept might be too abstract.

The *Street Projection* had the highest rate of participants entering the road before either one AV or both AVs fully stopped. Therefore, one could argue that because it communicates where exactly the vehicle is going to stop, the *Street Projection* promotes early crossing behavior while the *Baseline* condition (no eHMI used) unintentionally encourages 'safe' crossing behavior. The results regarding *Understanding*, *Usefulness* and *Satisfying* support this theory. The *Progress Bar* and *Street Projection* were rated significantly more positive compared to the *Baseline* in both analyses. As they are the only concepts communicating where the AV will stop, it shows that this information has a significant impact.

The fact that the *Pedestrian Symbol* was rated best on average and its positive ratings in terms of *Usefulness*, *Trust*, *Understanding*, and *Satisfying* in both of our analyses, the relatively high rate of "safe" *Crossing Behavior* and the positive open feedback might indicate that eHMIs don't have to be complex or abstract to work in a mixed-traffic scenario. They should rather be simple, established concepts where pedestrians might have encountered scalability in the past - like crossing a road at a pedestrian traffic light. It also implies that eHMIs should address a pedestrian ('You can cross') rather than only communicate their own intent/vehicle state. However, suggesting whether or not it is safe to cross is a liability in a real-world scenario, especially in a failure case [46].

The *Effect of Yielding Target* analysis showed that participants considered any eHMI more useful than the *Baseline*, whereas only the *Progress Bar*, *Pedestrian Symbol*, and *Street Projection* were significantly more useful than the *Baseline* in the *Mere Presence Effect* analysis. This shows the importance of research in the area of scalability regarding pedestrians, as only 3 out of 8 eHMI concepts were deemed more useful when more than one pedestrian was involved compared to no additional pedestrians, which is often the eHMI evaluation scenario [18].

5.4 How Should Automated Vehicles Communicate in a Complex Environment?

Even though only one 'dangerous' collision occurred in our filtered analysis, we emphasize that eHMI concepts should consider other vehicles, especially in a mixed-traffic scenario. An eHMI like the *Pedestrian Symbol* should never communicate that it is safe to cross if there are still other vehicles between the AV and the pedestrian.

Overall, it needs to be discussed what is 'desired' crossing behavior and whether or not an eHMI encourages it, especially under the scalability aspect of multiple lanes and vehicles. We argue that an eHMI should likely not encourage a pedestrian to enter the road before both vehicles stop. Although they might not be in immediate danger when waiting in a lane where an AV yielded for them, eHMIs should not motivate pedestrians to enter a road with ongoing traffic. Previous studies [7, 25, 27, 72] measured the 'willingness to cross' in a scenario with one lane and one vehicle. Here, an early lane entry - if there is enough time for the pedestrian to cross before a collision - only has a limited effect on the traffic because there are no other lanes. However, in a mixed-traffic scenario with multiple lanes and multiple vehicles in those lanes, a pedestrian entering a lane early and holding up traffic could cause accidents - especially if the pedestrian misinterpreted the AV's yielding intentions and forced it to stop earlier than anticipated. A vulnerable road user should certainly not be standing on the road at that point. While our *Multilane Bumper Text* solution did not perform particularly well, we urge designers to investigate further the concept of a 'representative vehicle'/'omniscient narrator' [15] or V2V-based solutions in general. For example, a variation of the original concept of the *Pedestrian Symbol* by Clamann et al. [7] could be used. It displays a crossed-out pedestrian symbol while slowing down. However, instead of turning into a pedestrian symbol once it stopped, it would only do that once traffic on both lanes comes to a full stop. This example could even be extended into a 2-part eHMI that communicates where the AV will stop. However, adding too much information could lead to overload and, thus, to misinterpretation and early lane entry again.

5.5 Limitations

Our study did not feature audio-based eHMI concepts and did not incorporate sound as we could not sufficiently control this variable. Nonetheless, it is important to focus on non-visual concepts in an attempt to find more inclusive eHMI concepts for the safety of pedestrians with vision impairments [15, 18].

As mostly younger participants (on average 30 years old) took part, it is unclear whether this work's findings are transferable to other age groups.

Finally, the setting with four lanes and heavy traffic represents an extreme case to cross the road. While, in general, this might not be advisable, we opted for this extreme case because of the symmetrical layout (having a three-lane road would have induced unbalanced traffic). To reduce the effect of this extreme case, participants only walked to the middle waiting strip. In this study, we evaluated **four** scenarios crossed with **nine** eHMIs derived from the literature, totaling **36** conditions.

5.6 Future Work

We highlight that future work should consider additional scenarios referring to scalability. Colley and Rukzio [14] provide a good starting point for relevant situational parameters to consider. Additionally, there are also numerous other parameters such as culture and faith that impact crossing decisions [17].

6 CONCLUSION

Due to the lack of research regarding the scalability of eHMIs [18], we developed the browser game 'Cross the Road' in which participants take control over a pedestrian tasked to cross a street in a mixed-traffic environment. In our study (N=46), we compared nine eHMI concepts (*Bumper Light*, *Bumper Text*, *Multilane Bumper Text*, *Pedestrian Symbol*, *Smiling Car*, *Situational Awareness Indicator*, *Progress Bar*, *Street Projection*, and one *base*) regarding scalability by combining them with four pedestrian conditions: Participants would either cross the road alone or with two other pedestrians where the AVs would yield for one of the three pedestrians. Our results showed that the *Pedestrian Symbol*, *Street Projection*, and *Progress Bar* performed best. It shows that communicating whether or not it is safe to cross either directly by telling the pedestrian (*Pedestrian Symbol*) or indirectly by communicating where the vehicle is going to stop (*Street Projection*, *Progress Bar*) was perceived best.

We argue that eHMIs should not promote early lane-entering behavior and should, thus, only communicate that it is safe to enter the road and cross once traffic on all lanes has stopped. Designers should keep eHMIs simple and familiar while keeping multiple lanes in mind. Our work helps to introduce AVs safely into general traffic.

ACKNOWLEDGMENTS

The authors thank all study participants. This work was conducted within the project 'Interaction between automated vehicles and vulnerable road users' (Intuitiver), funded by the Ministry of Science, Research and Arts of the State of Baden-Württemberg.

A COLLISION PLOTS

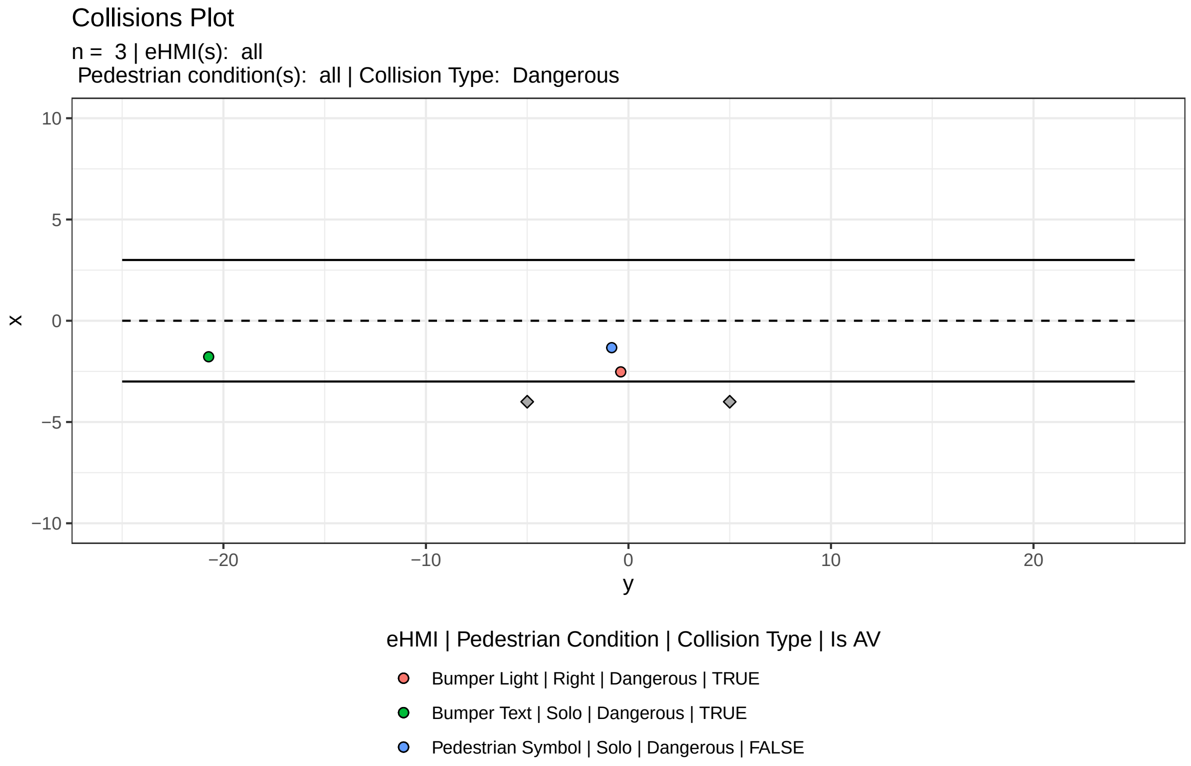


Fig. 7. Locations of the dangerous collisions for the filtered game log data set. The dangerous collision with an AV equipped with the *Bumper Text* can be ruled as an outlier as the participant in the trial walked to the left (most likely to test the boundaries of the game world). The collision with the AV during the *Bumper Light* x *Right* condition indicates that the participant misinterpreted the AV's yielding intention. The collision with a Non-AV during the *Pedestrian Symbol* x *Solo* condition could've occurred because the participant only focused on the AV's eHMI and didn't pay attention to the vehicle right in front of them.

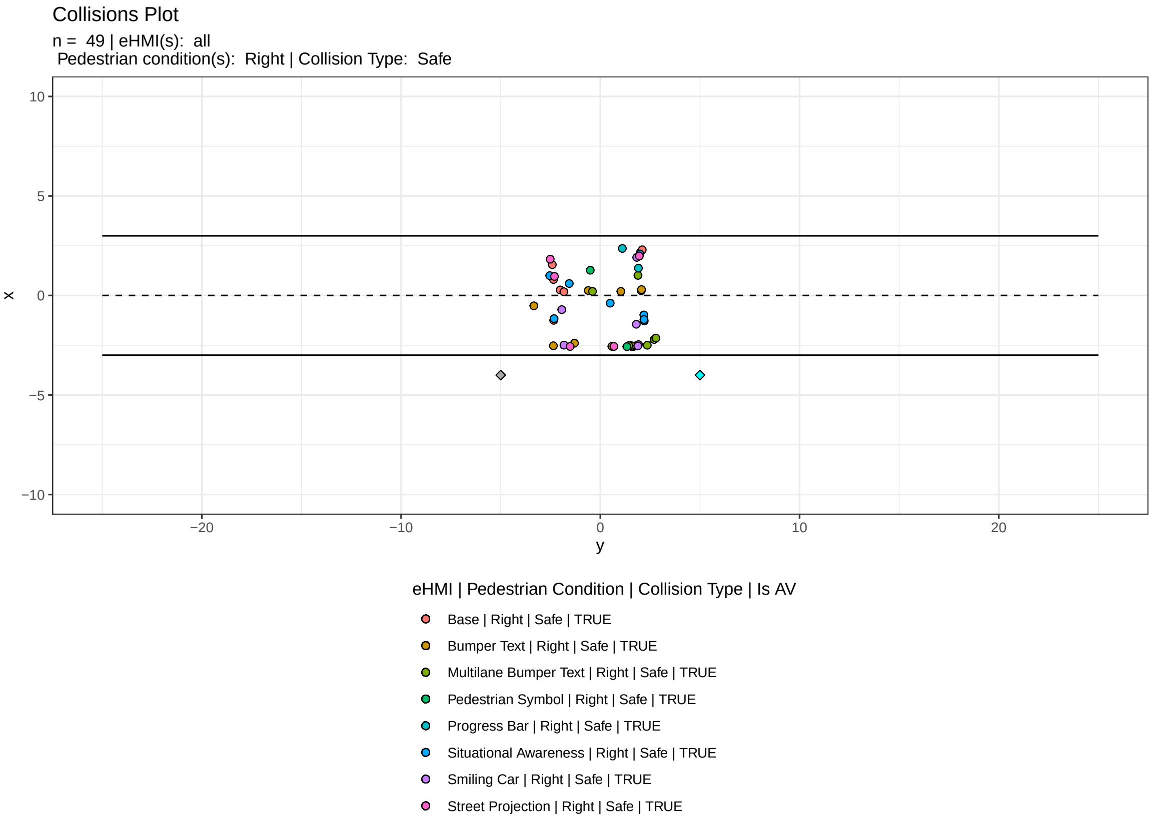


Fig. 8. Locations of *Safe* collisions with AVs during the *Right* pedestrian condition using the filtered game log dataset. Most collisions happen on the right, around the area where participants have to walk around the vehicle as it yields for the participant to their right. The collisions on the left occurred if the participant forced the AV to yield earlier.

REFERENCES

[1] Claudia Ackermann, Matthias Beggiato, Sarah Schubert, and Josef F Krems. 2019. An experimental study to investigate design and assessment criteria: what is important for communication between pedestrians and automated vehicles? *Applied ergonomics* 75 (2019), 272–282.

[2] 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* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376197

[3] Ashratuz Zavin Asha, Christopher Smith, Georgina Freeman, Sean Crump, Sowmya Somanath, Lora Oehlberg, and Ehud Sharlin. 2021. Co-Designing Interactions between Pedestrians in Wheelchairs and Autonomous Vehicles. In *Designing Interactive Systems Conference 2021* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 339–351. https://doi.org/10.1145/3461778.3462068

[4] AutoMotoTV. 2015. Mercedes-Benz F 015 Luxury in Motion - Interaction Zebra Crossing | AutoMotoTV. https://www.youtube.com/watch?v=fvtlobbMENo

[5] Magnus Carlsson and Per Nilsson. 2016. The smiling car - self driving car that sees you. https://semcon.com/smilingcar/

[6] Vicky Charisi, Azra Habibovic, Jonas Andersson, Jamy Li, and Vanessa Evers. 2017. Children's Views on Identification and Intention Communication of Self-Driving Vehicles. In *Proceedings of the 2017 Conference on Interaction Design and Children* (Stanford, California,

USA) *(IDC '17)*. Association for Computing Machinery, New York, NY, USA, 399–404. https://doi.org/10.1145/3078072.3084300

[7] Michael Clamann, Miles Aubert, and Mary L. Cummings. 2017. Evaluation of Vehicle-to-Pedestrian Communication Displays for Autonomous Vehicles. (2017), 13 pages.

[8] Claire D. Coles, Dorothy C. Strickland, Lynne Padgett, and Lynnae Bellmoff. 2007. Games that "work": Using computer games to teach alcohol-affected children about fire and street safety. *Research in Developmental Disabilities* 28, 5 (2007), 518–530. https://doi.org/10.1016/j.ridd.2006.07.001

[9] Mark Colley, Elvedin Bajrovic, and Enrico Rukzio. 2022. Effects of Pedestrian Behavior, Time Pressure, and Repeated Exposure on Crossing Decisions in Front of Automated Vehicles Equipped with External Communication. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 367, 11 pages. https://doi.org/10.1145/3491102.3517571

[10] Mark Colley, Jan Henry Belz, and Enrico Rukzio. 2021. Investigating the Effects of Feedback Communication of Autonomous Vehicles. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. Association for Computing Machinery, New York, NY, USA, 263–273. https://doi.org/10.1145/3409118.3475133

[11] Mark Colley, Christian Hummler, and Enrico Rukzio. 2022. Effects of mode distinction, user visibility, and vehicle appearance on mode confusion when interacting with highly automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour* 89 (2022), 303–316. https://doi.org/10.1016/j.trf.2022.06.020

[12] Mark Colley, Surong Li, and Enrico Rukzio. 2021. Increasing Pedestrian Safety Using External Communication of Autonomous Vehicles for Signalling Hazards. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction* (Toulouse & Virtual, France) *(MobileHCI '21)*. Association for Computing Machinery, New York, NY, USA, Article 20, 10 pages. https://doi.org/10.1145/3447526.3472024

[13] Mark Colley, Stefanos Can Mytilineos, Marcel Walch, Jan Gugenheimer, and Enrico Rukzio. 2020. Evaluating Highly Automated Trucks as Signaling Lights. 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/10.1145/3409120.3410647

[14] Mark Colley and Enrico Rukzio. 2020. A Design Space for External Communication of Autonomous Vehicles. 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/10.1145/3409120.3410646

[15] Mark Colley, Marcel Walch, Jan Gugenheimer, Ali Askari, and Enrico Rukzio. 2020. Towards Inclusive External Communication of Autonomous Vehicles for Pedestrians with Vision Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, Hawaii USA) *(CHI '20)*. ACM, Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3313831.3376472

[16] Mark Colley, Marcel Walch, Jan Gugenheimer, and Enrico Rukzio. 2019. Including People with Impairments from the Start: External Communication of Autonomous Vehicles. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 307–314. https://doi.org/10.1145/3349263.3351521

[17] Mark Colley, Marcel Walch, and Enrico Rukzio. 2019. For a Better (Simulated) World: Considerations for VR in External Communication Research. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 442–449. https://doi.org/10.1145/3349263.3351523

[18] Mark Colley, Marcel Walch, and Enrico Rukzio. 2020. Unveiling the Lack of Scalability in Research on External Communication of Autonomous Vehicles. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, Hawaii USA) *(CHI '20)*. ACM, Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3334480.3382865

[19] Mark Colley, Bastian Wankmüller, Tim Mend, Thomas Väth, Enrico Rukzio, and Jan Gugenheimer. 2022. User gesticulation inside an automated vehicle with external communication can cause confusion in pedestrians and a lower willingness to cross. *Transportation Research Part F: Traffic Psychology and Behaviour* 87 (2022), 120–137. https://doi.org/10.1016/j.trf.2022.03.011

[20] Shuchisnigdha Deb, Daniel W. Carruth, Muztaba Fuad, Laura M. Stanley, and Darren Frey. 2020. Comparison of Child and Adult Pedestrian Perspectives of External Features on Autonomous Vehicles Using Virtual Reality Experiment. In *Advances in Human Factors of Transportation*, Neville Stanton (Ed.). Springer International Publishing, Cham, 145–156.

[21] Debargha Dey, Azra Habibovic, Melanie Berger, Devanshi Bansal, Raymond H. Cuijpers, and Marieke Martens. 2022. Investigating the Need for Explicit Communication of Non-Yielding Intent through a Slow-Pulsing Light Band (SPLB) EHMI in AV-Pedestrian Interaction. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Seoul, Republic of Korea) *(AutomotiveUI '22)*. Association for Computing Machinery, New York, NY, USA, 307–318. https://doi.org/10.1145/3543174.3546086

[22] 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.

[23] 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. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376325

[24] Debargha Dey, Kai Holländer, Melanie Berger, Berry Eggen, Marieke Martens, Bastian Pfleging, and Jacques Terken. 2020. Distance-Dependent EHMIs for the Interaction Between Automated Vehicles and Pedestrians. 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, 192–204. https://doi.org/10.1145/3409120.3410642

[25] Debargha Dey, Kai Holländer, Melanie Berger, Berry Eggen, Marieke Martens, Bastian Pfleging, and Jacques Terken. 2020. Distance-Dependent EHMIs for the Interaction Between Automated Vehicles and Pedestrians. 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, 192–204. https://doi.org/10.1145/3409120.3410642

[26] Debargha Dey, Marieke Martens, Chao Wang, Felix Ros, and Jacques Terken. 2018. Interface Concepts for Intent Communication from Autonomous Vehicles to Vulnerable Road Users. In *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Toronto, ON, Canada) *(AutomotiveUI '18)*. Association for Computing Machinery, New York, NY, USA, 82–86. https://doi.org/10.1145/3239092.3265946

[27] Debargha Dey, Arjen van Vastenhoven, Raymond H. Cuijpers, Marieke Martens, and Bastian Pfleging. 2021. *Towards Scalable EHMIs: Designing for AV-VRU Communication Beyond One Pedestrian.* Association for Computing Machinery, New York, NY, USA, 274–286. https://doi.org/10.1145/3409118.3475129

[28] Debargha Dey, Francesco Walker, Marieke Martens, and Jacques Terken. 2019. Gaze Patterns in Pedestrian Interaction with Vehicles: Towards Effective Design of External Human-Machine Interfaces for Automated Vehicles. 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, 369–378. https://doi.org/10.1145/3342197.3344523

[29] André Dietrich, Jan-Henning Willrodt, Karolin Wagner, and Klaus Bengler. 2018. Projection-Based External Human Machine Interfaces – Enabling Interaction between Automated Vehicles and Pedestrians. In *Proceedings of the Driving Simulation Conference 2018 Europe VR* (2018-09-05), Andras Kemeny, Florent Colombet, Frédéric Merienne, and Stéphane Espié (Eds.). Driving Simulation Association, Antibes, France, 43–50.

[30] Weihua Dong, Hua Liao, Bing Liu, Zhicheng Zhan, Huiping Liu, Liqiu Meng, and Yu Liu. 2020. Comparing pedestrians' gaze behavior in desktop and in real environments. *Cartography and Geographic Information Science* 47, 5 (2020), 432–451.

[31] 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

[32] 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.

[33] Daniel J Fagnant and Kara Kockelman. 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice* 77 (2015), 167–181.

[34] 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.

[35] Evelyn Florentine, Mark Adam Ang, Scott Drew Pendleton, Hans Andersen, and Marcelo H. Ang. 2016. Pedestrian Notification Methods in Autonomous Vehicles for Multi-Class Mobility-on-Demand Service. In *Proceedings of the Fourth International Conference on Human Agent Interaction* (Biopolis, Singapore) *(HAI '16)*. Association for Computing Machinery, New York, NY, USA, 387–392. https://doi.org/10.1145/2974804.2974833

[36] Tanja Fuest, Elisabeth Schmidt, and Klaus Bengler. 2020. Comparison of Methods to Evaluate the Influence of an Automated Vehicle's Driving Behavior on Pedestrians: Wizard of Oz, Virtual Reality, and Video. *Information* 11, 6 (2020), 22 pages. https://doi.org/10.3390/info11060291

[37] 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

[38] Mathias Haimerl, Mark Colley, and Andreas Riener. 2022. Evaluation of Common External Communication Concepts of Automated Vehicles for People With Intellectual Disabilities. In *Proc. ACM Hum.-Comput. Interact. (MobileHCI '22)*. Association for Computing Machinery, New York, NY, USA, 13 pages. https://doi.org/10.1145/3546717

[39] Mathias Haimerl, Mark Colley, and Andreas Riener. 2022. Evaluation of Common External Communication Concepts of Automated Vehicles for People With Intellectual Disabilities. *Proc. ACM Hum.-Comput. Interact.* 6, MHCI, Article 182 (sep 2022), 19 pages. https://doi.org/10.1145/3546717

[40] 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.

[41] Tobias Hecht, Anna Feldhütter, Kathrin Draeger, and Klaus Bengler. 2019. What do you do? An analysis of non-driving related activities during a 60 minutes conditionally automated highway drive. In *International Conference on Human Interaction and Emerging Technologies.*

Springer, Springer, 28–34.

[42] 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*. Association for Computing Machinery, New York, NY, USA, Article 158, 13 pages. https://doi.org/10.1145/3411764.3445480

[43] Kai Holländer, Andy Krüger, and Andreas Butz. 2020. Save the Smombies: App-Assisted Street Crossing. 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 22, 11 pages. https://doi.org/10.1145/3379503.3403547

[44] Kai Holländer, Luca Schellenberg, Changkun Ou, and Andreas Butz. 2020. All Fun and Games: Obtaining Critical Pedestrian Behavior Data from an Online Simulation. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3382797

[45] Ming Hou, Karthik Mahadevan, Sowmya Somanath, Ehud Sharlin, and Lora Oehlberg. 2020. Autonomous Vehicle-Cyclist Interaction: Peril and Promise. 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, 1–12. https://doi.org/10.1145/3313831.3376884

[46] Anees Ahamed Kaleefathullah, Natasha Merat, Yee Mun Lee, Yke Eisma, Ruth Madigan, J Garcia, and Joost de Winter. 2020. External Human-Machine Interfaces can be Misleading: An examination of trust development and misuse in a CAVE-based pedestrian simulation environment. *Human Factors* 0, 0 (09 2020). https://doi.org/10.1177/0018720820970751 arXiv:https://doi.org/10.1177/0018720820970751 PMID: 33242999.

[47] 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.

[48] Mirjam Lanzer, Ina Koniakowsky, Mark Colley, and Martin Baumann. 2023. Interaction Effects of Pedestrian Behavior, Smartphone Distraction and External Communication of Automated Vehicles on Crossing and Gaze Behavior. (2023). https://doi.org/10.1145/3544548.3581303

[49] 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, et al. 2021. Road users rarely use explicit communication when interacting in today's traffic: implications for automated vehicles. *Cognition, Technology & Work* 23 (2021), 367–380.

[50] Yee Mun Lee, Ruth Madigan, Chinebuli Uzondu, Jorge Garcia, Richard Romano, Gustav Markkula, and Natasha Merat. 2022. Learning to interpret novel eHMI: The effect of vehicle kinematics and eHMI familiarity on pedestrian' crossing behavior. *Journal of Safety Research* 80 (2022), 270–280. https://doi.org/10.1016/j.jsr.2021.12.010

[51] Letitia Lew, Truc Nguyen, Solomon Messing, and Sean Westwood. 2011. Of Course I Wouldn't Do That in Real Life: Advancing the Arguments for Increasing Realism in HCI Experiments. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI EA '11)*. Association for Computing Machinery, New York, NY, USA, 419–428. https://doi.org/10.1145/1979742.1979621

[52] 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* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 262–274. https://doi.org/10.1145/3342197.3344544

[53] Lumineq. 2022. Glass displays for improved automotive safety. https://www.lumineq.com/applications/automotive

[54] Victor Malmsten Lundgren, Azra Habibovic, Jonas Andersson, Tobias Lagström, Maria Nilsson, Anna Sirkka, Johan Fagerlönn, Rikard Fredriksson, Claes Edgren, Stas Krupenia, and Dennis Saluäär. 2017. Will There Be New Communication Needs When Introducing Automated Vehicles to the Urban Context?. In *Advances in Human Aspects of Transportation*, Neville A. Stanton, Steven Landry, Giuseppe Di Bucchianico, and Andrea Vallicelli (Eds.). Springer International Publishing, Cham, 485–497.

[55] Haiko Lüpsen. 2021. R-Funktionen zur Varianzanalyse. http://www.uni-koeln.de/~luepsen/R/. [Online; accessed 25-SEPTEMBER-2021].

[56] Karthik Mahadevan, Elaheh Sanoubari, Sowmya Somanath, James E. Young, and Ehud Sharlin. 2019. AV-Pedestrian Interaction Design Using a Pedestrian Mixed Traffic Simulator. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19)*. Association for Computing Machinery, New York, NY, USA, 475–486. https://doi.org/10.1145/3322276.3322328

[57] Karthik Mahadevan, Sowmya Somanath, and Ehud Sharlin. 2018. Communicating Awareness and Intent in Autonomous Vehicle-Pedestrian Interaction. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3174003

[58] Alexander G. Mirnig, Magdalena Gärtner, Vivien Wallner, Michael Gafert, Hanna Braun, Peter Fröhlich, Stefan Suette, Jakub Sypniewski, Alexander Meschtscherjakov, and Manfred Tscheligi. 2021. Stop or Go? Let Me Know! A Field Study on Visual External Communication for Automated Shuttles. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) *(AutomotiveUI '21)*. Association for Computing Machinery, New York, NY, USA, 287–295. https://doi.org/10.1145/3409118.3475131

[59] Dylan Moore, Rebecca Currano, and David Sirkin. 2020. Sound Decisions: How Synthetic Motor Sounds Improve Autonomous Vehicle-Pedestrian Interactions. 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, 94–103. https://doi.org/10.1145/3409120.3410667

[60] Trung Thanh Nguyen, Kai Holländer, Marius Hoggenmueller, Callum Parker, and Martin Tomitsch. 2019. Designing for Projection-Based Communication between Autonomous Vehicles and Pedestrians. 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, 284–294. https://doi.org/10.1145/3342197.3344543

[61] Bastian Pfleging, Maurice Rang, and Nora Broy. 2016. Investigating User Needs for Non-Driving-Related Activities during Automated Driving. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia* (Rovaniemi, Finland) *(MUM '16)*. Association for Computing Machinery, New York, NY, USA, 91–99. https://doi.org/10.1145/3012709.3012735

[62] Amir Rasouli, Iuliia Kotseruba, and John K Tsotsos. 2017. Understanding pedestrian behavior in complex traffic scenes. *IEEE Transactions on Intelligent Vehicles* 3, 1 (2017), 61–70.

[63] Jaguar Land Rover. 2018. The Virtual Eyes have it. https://www.jaguarlandrover.com/2018/virtual-eyes-have-it

[64] Semcon. 2016. The Smiling Car. https://www.youtube.com/watch?v=INqWGr4dfnU

[65] Andreas Sieß, Kathleen Hübel, Daniel Hepperle, Andreas Dronov, Christian Hufnagel, Julia Aktun, and Matthias Wölfel. 2015. Hybrid City Lighting-Improving Pedestrians' Safety through Proactive Street Lighting. In *2015 International Conference on Cyberworlds (CW)*. IEEE, New York, NY, USA, 46–49.

[66] Helena Strömberg, Ingrid Pettersson, Jonas Andersson, Annie Rydström, Debargha Dey, Maria Klingegård, and Jodi Forlizzi. 2018. Designing for social experiences with and within autonomous vehicles – exploring methodological directions. *Design Science* 4 (2018), e13. https://doi.org/10.1017/dsj.2018.9

[67] SAE Taxonomy. 2014. *Definitions for terms related to on-road motor vehicle automated driving systems*. Technical Report. Technical report, SAE International.

[68] Unity Technologies. 2022. *Unity*. Unity Technologies. https://unity.com/

[69] 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.

[70] Luis von Ahn and Laura Dabbish. 2004. Labeling Images with a Computer Game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vienna, Austria) *(CHI '04)*. Association for Computing Machinery, New York, NY, USA, 319–326. https://doi.org/10.1145/985692.985733

[71] 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.

[72] Marc Wilbrink, Manja Nuttelmann, and Michael Oehl. 2021. Scaling up Automated Vehicles' EHMI Communication Designs to Interactions with Multiple Pedestrians – Putting EHMIs to the Test. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. Association for Computing Machinery, New York, NY, USA, 119–122. https://doi.org/10.1145/3473682.3480277

[73] Xingchen Zhang, Panagiotis Angeloudis, and Yiannis Demiris. 2022. ST CrossingPose: A Spatial-Temporal Graph Convolutional Network for Skeleton-Based Pedestrian Crossing Intention Prediction. *IEEE Transactions on Intelligent Transportation Systems* 23, 11 (2022), 20773–20782.

[74] Xiangling Zhuang and Changxu Wu. 2011. Pedestrians' crossing behaviors and safety at unmarked roadway in China. *Accident Analysis & Prevention* 43, 6 (2011), 1927–1936. https://doi.org/10.1016/j.aap.2011.05.005

[75] Xiangling Zhuang and Changxu Wu. 2011. Pedestrians' crossing behaviors and safety at unmarked roadway in China. *Accident Analysis & Prevention* 43, 6 (2011), 1927–1936. https://doi.org/10.1016/j.aap.2011.05.005

[76] Raphael Zimmermann and Reto Wettach. 2017. First Step into Visceral Interaction with Autonomous Vehicles. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Oldenburg, Germany) *(AutomotiveUI '17)*. Association for Computing Machinery, New York, NY, USA, 58–64. https://doi.org/10.1145/3122986.3122988