Augmented reality for supporting the interaction between pedestrians and automated vehicles: An experimental outdoor study

18 October 2023

Thomas K. Aleva*, Wilbert Tabone*, Dimitra Dodou, Joost C.F. de Winter

Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology

* Joint first authors

Corresponding author: j.c.f.dewinter@tudelft.nl

Abstract

Communication from automated vehicles (AVs) to pedestrians using augmented reality (AR) could positively contribute to traffic safety. However, previous research was mainly conducted through questionnaires or experiments in virtual instead of real environments. In this study, 28 participants conducted trials outdoors with an approaching AV and were supported by four different AR interfaces. The AR experience was created by having participants wear a Varjo XR-3 headset with see-through functionality, with the AV and AR elements virtually overlaid onto the real environment. The AR interfaces were vehicle-locked (*Planes on vehicle*), world-locked (*Fixed pedestrian lights*; *Virtual fence*), or head-locked (*Pedestrian lights HUD*). Participants had to hold down a button as long as they felt safe to cross, and their opinions were obtained through rating scales, interviews, and a questionnaire. The results showed that AR interfaces were more preferred and effective than no AR interface, evidenced by higher perceived safety to cross for yielding AVs. The *fixed pedestrian lights* scored lower than the other interfaces, presumably due to low saliency and the fact that participants had to visually identify both the AR interface and the AV. In conclusion, in our outdoor assessment of AR interfaces for AV-pedestrian interactions, AR was preferred over no interface, but preference depended on design elements like placement, salience, and visual attention requirements. Looking ahead, as AR gains more prominence, the need for human-subject studies in real-world settings becomes crucial. This study offers insights into the potential benefits and challenges of implementing AR in real-world scenarios.

Keywords: Augmented reality; Pedestrian safety; Anchoring; See-through AR

Introduction

Every year, 300,000 pedestrian deaths occur worldwide, accounting for 23% of all road fatalities (World Health Organization, 2018). The United Nations has set a target within its Sustainable Development Goals to halve the number of traffic deaths by 2030 (United Nations, 2020). In this resolution (A/RES/74/299), it was also noted that "continuous progress of automotive and digital technologies could improve road safety, including through the progressive development of highly and fully automated vehicles in road traffic" (United Nations, 2020, p. 4).

One possibility to improve the safety of pedestrians is to equip automated vehicles (AVs) with better sensors and intelligence so that they can respond earlier, in order to reduce collisions with vulnerable road users (Jungmann et al., 2020; Paiva et al., 2021). Another possibility concerns the introduction of AV-to-pedestrian communication technologies. Previous research has shown that with displays on the outside of the AV, called external human-machine interfaces or eHMIs, pedestrians can make more effective road-crossing decisions (Bazilinskyy et al., 2021; Bindschädel et al., 2022; Dey et al., 2020). However, eHMIs have certain disadvantages in that they typically cannot address an individual pedestrian (Colley et al., 2020; Tran et al., 2023) and that they can be difficult to perceive in some cases, for example due to a temporary occlusion by another object (Dey et al., 2022; Troel-Madec et al., 2019).

Recently, AV-to-pedestrian communication via wearable devices has been proposed (Gelbal et al., 2023; Hasan & Hasan, 2022; Lakhdhir et al., 2023). More specifically, augmented reality (AR), defined as a technology that overlays virtual information onto the real world, appears to be an increasingly promising solution for communicating with pedestrians (Tabone et al., 2021b). On the one hand, the use of AR might sound unusual and undesirable; after all, it may be questioned whether pedestrians would want to rely on expensive technology in order to move safely through traffic (e.g., Berge et al., 2022; Tabone et al., 2021a). At the same time, it is recognized that AR in other forms, such as head-up displays (HUDs) in cars (Transparency Market Research, 2023) or the use of visual overlays in Google Maps on mobile phones (Google, 2020), is already commonplace. Especially in the context of developments such as the introduction of the Apple Vision Pro (Apple, 2023) and the types of wearables that are already widely used, like smartwatches, it is conceivable that in the future, pedestrians will receive verbal and visual assistance through AR.

Although AR is a potentially promising technology for pedestrians, still little research exists on this topic. One of the problems is that AR is currently challenging to implement and that only the most recent headsets are capable of creating a compelling and user-friendly experience (e.g., Microsoft Hololens 2, Varjo XR 3, Magic Leap 2). While exceptions exist (e.g., Kang et al., 2023; Zhang et al., 2023), much previous AR research for pedestrians has limitations: some studies focus solely on the orienting phase without any form of human-subject evaluation (Tabone et al., 2021b; Tong et al., 2021), others use online questionnaires with photos or videos (Hesenius et al., 2018; Tabone et al., 2023a; Wilbrink et al., 2023), and still others test AR concepts in immersive virtual reality (VR) environments (Malik et al., 2023; Pratticò et al., 2021; Tabone et al., 2023b; Tran et al., 2022). A critical note that should be made in this last point is that—although valid results can be obtained in VR—this is strictly speaking not AR; after all, adding virtual displays to an already virtual environment remains VR, not AR. Such research does not require an AR headset or anchoring techniques, and the AR design is therefore only considered in a hypothetical form.

In this study, we share the findings from an outdoor experiment involving participants using an AR headset. We introduced a paradigm that incorporates AR, while also simulating certain real-world elements. Specifically, in addition to the virtual interfaces which are inherent to AR, we also simulated the approach of an AV to ensure consistent experimental conditions. Our approach serves as a stepping stone to a full AR experience, where pedestrians can walk around untethered while the AR headset wirelessly communicates with an actual AV. This use of AR, where virtual road users coexist with real humans in an outdoor space, has been applied by several others before, though not in the context of AV-pedestrian communication (Kamalasanan et al., 2022; Maruhn et al., 2020). Our methodology also aligns with other realistic simulation methods, such as displaying virtual road users in a real car (Bokc et al., 2007; Butenuth et al., 2017; Feng et al., 2018; Hussain et al., 2013; Sheridan, 2016), a technique that supersedes driving simulators which inherently offer limited visual and motion cues. In our approach, while the environment remains the real world, it is enhanced with simulated experimental stimuli.

In the current experiment, four different AR interfaces and a no-AR-interface baseline condition were compared among human participants. The focus was on whether these AR interfaces made participants feel safe to cross, as well as subjective qualities including whether the AR interface was found to be intuitive and was accepted. The interfaces used in this study were adopted from three prior works: an AR concept design study (Tabone et al., 2021b), an online questionnaire with video clips (Tabone et al., 2023a), and a CAVE-based virtual simulator study (Tabone et al., 2023b). Our goal was to investigate whether the results we found in the current experiment correspond with the earlier research among human participants (Tabone et al., 2021b, 2023a).

The candidate AR interfaces differed in terms of their anchoring techniques: the AR interface was either AV-locked, world-locked, or head-locked, three methods that bring fundamentally different demands regarding the user's visual attention distribution (Lebeck et al., 2017; Lingam et al., 2023; Peereboom et al., 2023; Tabone et al., 2023b). A vehicle-locked AR interface entails, just like an eHMI, that implicit communication in the form of the speed of and distance to the AV and explicit AR cues are congruent in time and place. A head-locked AR interface, on the other hand, is always visible but requires that the AV is looked at to confirm the cues of the AR interface. Finally, a world-locked interface requires attention to be distributed between the AV and the AR interface, with the possible advantage that the AR interface can be presented at a fixed and familiar location (Lingam et al., 2023). In addition to comparing different AR interfaces with no AR interface, one of the objectives of this research was to document, accompanied by Unity source code, how AR research in the outdoor environment can be conducted.

Methods

Participants

Participants were recruited through personal networks. No financial reimbursement or other incentives were offered. A total of 28 people, of which 23 males and 5 females, aged between 19 and 59 (M = 27.2, SD = 9.2), participated. Nineteen of the participants were students. Twenty-six participants indicated being Dutch, one Chinese, and the nationality of one participant was unknown. Twenty-six participants held a driver's licence, for an average of 8.0 (SD = 7.3) years. A total of 17 participants had used a VR headset before while 9 participants had used an AR headset before. Daily walking time was less than 15 min for 7 participants, 15 to 30 minutes for 12 participants, and more than 30 min for 9 participants. Cycling was the primary mode of transportation for most participants (22 out of 28). Each participant provided written informed consent. The experiment procedure was approved by the TU Delft Human Research Ethics Committee, approval no. 3054.

Materials

The experiment took place in a designated area of the Delft University of Technology campus, which was closed to traffic. It was conducted on an Alienware PC powered by an Intel® Core™i7-9700K CPU at 3.60 GHz, equipped with 16 GB RAM and a 16 GB Nvidia GeForce RTX 2080 Ti GPU. The AR software ran in Unity 2021.3.13f1 (Unity, 2022), combined with the Varjo XR Plugin (Varjo Developer, 2023). A custom script was used that allowed the experimenter to select the experiment conditions from within Unity.

AR was displayed by means of a Varjo XR-3 headset. The Varjo VR-3 provides a 90 Hz refresh rate and a 115° horizontal field of view. The focus area, of 27° × 27°, was rendered at 70 pixels per degree on a µOLED display, providing 1920 × 1920 pixels per eye. The peripheral area was rendered at about 30 pixels per degree on an LCD, producing 2880 × 2720 pixels per eye. The Varjo XR-3 presented a virtual AV and virtual AR interfaces while depicting the real world by means of video pass-through. A pole was used to route the cables from the PC to the participant (Figure 1).

Our initial intention for anchoring the virtual objects to the real world was to use object tracking and reference markers (Varjo Technologies, 2023). However, this proved to be unreliable, presumably due to the relatively empty environment and occasional sunlight reflections. Instead, tracking for the headset was achieved using two SteamVR Base Stations 2.0 (HTC VIVE, 2019), as depicted in Figure 1. Each time the hardware was set up, calibration was performed first. Our solution was to position the *AR Origin* according to the calibration position and orientation of the Varjo XR-3. The scale value of the *AR Origin* was set such that distances in the virtual world corresponded with distances in the real world. In Unity, an invisible *Ground place* was created, which functioned as a surface for the AV to drive on. The height of the *Main Camera* above the ground plane was 1.70 m.

A Logitech R400 wireless presenter was used as the remote control. The right arrow button was used by the participant as the button to indicate if the participant felt it was safe to cross. The receiver was connected via a USB port of the Alienware desktop.

The sound of the AV was transmitted through headphones that were plugged into the Varjo XR-3. The headphones did not include noise-cancelling, so surrounding real-world environment noise could still be heard.

Figure 1. Participant wearing the tethered Varjo XR-3 headset.

Participant Information and Task
Participants were provided with information about the procedure and tasks through a leaflet. They were informed that they would wear an AR headset, which uses cameras to capture the environment and can project virtual objects onto this real-world setting.

Participants were informed that they would be positioned at the side of a road and that their task was to determine whether they felt safe to cross the road. Their specific task was to press and hold the button when they felt it was safe to cross and to release it when they did not feel safe to cross. It was emphasised that participants should not physically cross the road.

The leaflet also mentioned that, at the start of each trial, participants would see a circle (hereafter named 'attention-attractor circle') either in front of them or to their left or right, which they were instructed to gaze at for a duration of one second to initiate the simulation. It was mentioned that a virtual vehicle would approach from the right, potentially stopping for the participant and possibly communicating this intent through various communication interfaces. Finally, it was mentioned that, after each trial, participants would be prompted with a question displayed in the AR environment, to which they were to provide a verbal response, and that this procedure would be repeated for a total of four different AR interfaces, plus a no-interface baseline condition. Participants were also informed that, following every block with a particular AR condition, the experimenter would ask open-ended questions about their experiences with the presented condition.

Automated Vehicle Behaviour
In each trial, a virtual AV approached from the right, displaying one of two behaviours: non-yielding or yielding. The vehicle model asset had dimensions of 4.95 m in length, 2.10 m in width, and 1.35 m in height (Final Form Studio, 2021). The AV's wheel movements were animated to match its speed. Audio was incorporated to produce a speed-sensitive engine sound, with a Doppler effect applied.

When the trial began, the virtual AV appeared approximately 45 m away from the participant, with the distance measured parallel to the road. Simultaneously, an attention-attractor stimulus in the form of a cyan circle ring appeared either in front of, to the left of, or to the right of the participant. After the participant looked at the circle for 1 s, the AV began driving at a speed of 30 km/h.

Once 3.1 s had elapsed from the moment that the AV began driving, it passed an invisible trigger, prompting the AR interface to appear. In the yielding condition, the AV began decelerating 3.9 s from the start of its movement and came to a complete stop at an elapsed time of 7.8 s. In the non-yielding condition, the AV maintained a speed of 30 km/h and passed the participant 6.5 s from the beginning of its movement. Figure 2 provides an aerial view of the roadway, illustrating the sequence of events: the AV starting, the appearance of the AR interface, the point where the AV began decelerating, and where it came to a full stop.

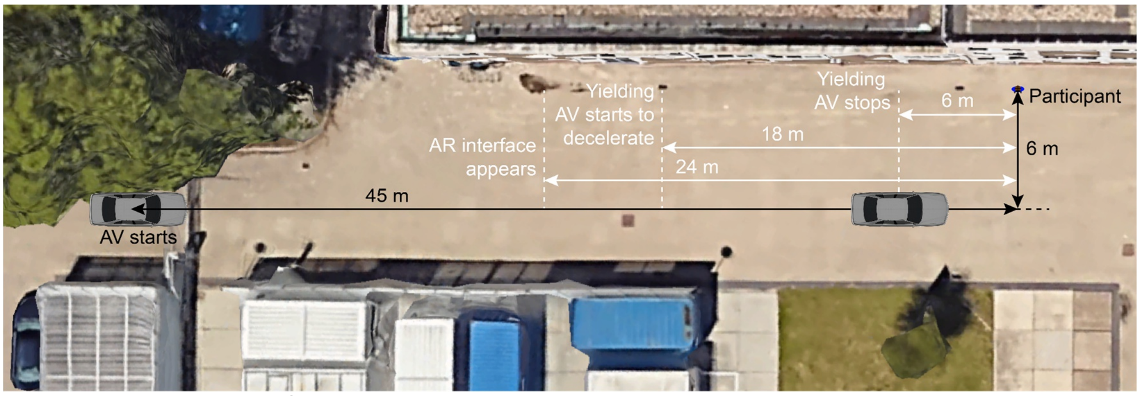

Figure 2. Top-down view of the experiment area.

Augmented Reality Designs

The AR interfaces were adopted from earlier research in which these interfaces had been designed (Tabone et al., 2021b), presented to a large sample of online respondents (Tabone et al., 2023a), and subjected to experimental testing in an immersive virtual simulator (Tabone et al., 2023b). The nine AR interfaces were divided into three anchoring methods: world-locked, head-locked, and vehicle-locked. Our previous experiments had already established which AR interfaces were effective and ineffective. Therefore, we focused on comparing different anchoring categories, an approach that was more concise than testing all conditions, making it suitable for outdoor experiments. Specifically, from each category, one was selected: *Virtual fence* (world-locked), *Pedestrian lights HUD* (head-locked), and *Planes on vehicle* (vehicle-locked). In addition, the *Fixed pedestrian lights* interface was selected. The reason for including this world-locked interface was to allow a comparison with the *Pedestrian lights HUD*. The four selected AR interfaces were as follows:

1. The *Virtual fence* consisted of a zebra crossing projected on the road combined with vertical semi-translucent walls surrounding the zebra (Figure 3). The interface was 3.0 m tall, 2.5 m wide and 7.5 m long. A semi-translucent gate positioned in front of the pedestrian opened in 1-s time after the interface appeared in the yielding condition, and remained closed in the non-yielding condition.

2. The *Fixed pedestrian lights* resembled existing pedestrian traffic lights (Figure 4). It remained stationary across the street from where the pedestrian was positioned. The interface was 2 m tall and positioned 15 m from the pedestrian. It consisted of a pole with a box (0.20 m wide, 0.38 m tall) on top displaying either a lit-up red icon of a standing pedestrian, or a green icon of a walking pedestrian.

3. The *Pedestrian lights HUD* was similar to the *Fixed pedestrian lights* interface (Figure 5). However, instead of being world-locked, the same box was anchored to the user’s head. It was placed at 1 m distance from the participant’s head and off-centre by 30 cm upwards and to the right. The interface was rotated around the vertical axis in order to face the user.

4. The *Planes on vehicle* was a vehicle-mapped interface consisting of a red plane with a stop-hand icon or a green plane with an icon depicting a pedestrian crossing a zebra crosswalk (Figure 6). The 2.1 wide and 1.58 m tall plane hovered above the front of the AV so that the bottom of the plane aligned with the front bumper; it was tilted by 54° to be positioned parallel with the AV's windshield.

In all four AR interfaces, colour was used to provide a redundant cue; for the non-yielding AVs this was pure red, and for yielding AVs this was pure green. The *Virtual fence* and *Planes on vehicle* were semi-transparent, to ensure that the AV or possibly other relevant objects remained visible to the participants.

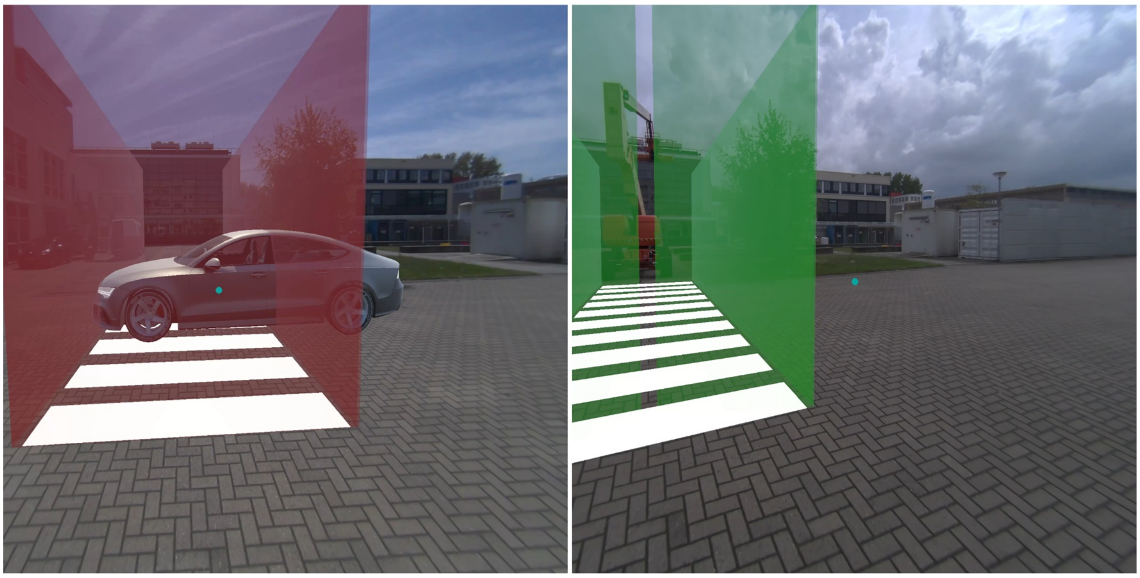

Figure 3. Virtual fence in the non-yielding and yielding condition.

Figure 4. Fixed pedestrian lights in the non-yielding and yielding condition.

Figure 5. Pedestrian lights HUD in the non-yielding and yielding condition.

Figure 6. Planes on vehicle in the non-yielding and yielding condition.

Experiment Design

The experiment was of a within-subject design. Each participant was exposed to the four AR interfaces and a no-interface baseline condition, two AV behaviours (yielding and non-yielding), and three attention-attractor locations (left, middle, right).

Each of the five interface conditions was presented in a separate block. Each block consisted of six trials: three with a yielding AV and three with a non-yielding AV. The three trials were conducted with the attention attractor at the left, right, or middle. This resulted in a total of 30 trials per participant (5 interface conditions × 2 yielding behaviours × 3 attention-attractor locations). The order of the five blocks, as well as the order of the six trials within each block, were counterbalanced using a Latin Square method.

The experiment lasted 45 to 60 minutes per participant, with the time spent wearing the headset amounting to approximately 30 minutes.

Questionnaires and Rating Scales

After signing the consent form, participants completed an intake questionnaire, designed in Qualtrics (Qualtrics XM, 2023), to gather demographic information. The intake questionnaire also included items about affinity for technology and items about whether participants had experienced

VR or AR headsets before. Some participants had completed the pre-experiment questionnaire before their scheduled experiment slot.

Next, participants were asked to read a leaflet describing a short description of the experiment. A complementary oral explanation of the experiment was provided where needed. Subsequently, participants put on the Varjo XR-3 headset, were handed the remote button, and a multi-point eye-tracker calibration was conducted. After each trial, at exactly 10 seconds after the AV had started moving, a statement appeared in front of the participant: "*This interface/situation was intuitive for signalling: 'Please do cross the road'.*" for the yielding condition, or "*This interface/situation was intuitive for signalling: 'Please do not cross the road'.*" for the non-yielding condition. Participants verbally indicated to what extent they agreed on a scale from *Fully disagree* (1) to *Fully agree* (7).

After each block of six trials with a particular AR condition, a semi-structured interview was conducted regarding interface design qualities, the timing of the interface appearance, the preference between yielding and non-yielding state, and the participant's wellbeing according to the misery scale (MISC; Bos, 2015).

After all five blocks, participants completed a post-experiment questionnaire in Qualtrics. This questionnaire contained items related to the AR experience and about AR interfaces in general. Participants were also asked to rank the five interface conditions in terms of their preference. Additionally, they were asked, for the four AR interfaces, to answer items regarding the intuitiveness of the green and red interfaces, convincingness of the green and red interface, iinterface trustworthiness, size (too small, too large), timing (too early, too late), clarity/understandability, and visual attractiveness, as well as a 9-item acceptance scale (Van der Laan et al., 1997), identical to Tabone et al. (2023a). The items used 7-point scales, except for the acceptance scale and adoption questions which use a 5-point semantic differential scale, and the ranking item. Open questions were also asked per interface condition to allow participants to justify their responses.

Data Recording and Analysis

The data was stored at a frequency of 50 Hz. Firstly, we determined per trial what percentage of the time participants kept the response button pressed. This was done from the moment the AR interface appeared until the vehicle came to a stop (yielding AVs) or passed (non-yielding AVs).

Additionally, from the post-experiment questionnaire, we determined a composite score per AR interface, identical to how it was done by Tabone et al. (2023b). The composite score was calculated from the participants' responses to 15 items, which included the 9-item acceptance scale and 6 other items (1. intuitiveness for non-yielding AVs, 2. intuitiveness for yielding AVs, 3. convincingness for non-yielding AVs, 4. convincingness for yielding AVs, 5. clarity/understandability, 6. attractiveness). The composite score was calculated by first appending the questionnaire results of the 28 participants and 4 AR conditions (resulting in a 112 × 15 matrix), and then applying a z-transform, followed by summing the 15 item scores, and again applying a z-transform. The reason for using a composite score is that the self-reports appeared to correlate strongly, and there was no evidence of multiple underlying constructs.

The preference rank, in which participants had to sort the five AR conditions from 1 (most preferred) to 5 (least preferred), was analysed as a separate item. Finally, the post-trial intuitiveness ratings were averaged over 3 trials per participant so that for each AR interface and participant, and for both yielding and non-yielding AV, an intuitiveness score was available.

The attention-attractor circles were present before the start of the trial, and the AR interface appeared 3.1 s later. Thus, participants had time to orient themselves before the AR interface appeared. The circles thus served the function of ensuring that participants did not continuously direct their attention towards where the AV appeared, resulting in more naturalistic behaviour compared to when no circles would be used. Since the circle locations did not have a fundamental influence on how the AR interfaces were responded to, the results for the three trials per AR interface and yielding condition were averaged.

In the statistical analyses, repeated-measures ANOVAs were used, with the AR condition as the independent variable. The findings are shown as means, complemented by 95% confidence intervals for within-subjects designs (Morey, 2008).

The interview results per AR interface and participant were summarised into short highlights by one of the authors. These highlights of all 28 participants were then analysed per AR condition using a thematic analysis, with an emphasis on whether the participants found the AR clear and whether they were satisfied with it.

Results

The 28 participants performed a total of 840 trials. 3 of 840 log files were unavailable or could not be used due to an experimenter error. For the post-experiment questionnaire, responses for the *Planes on vehicle* were unavailable for 1 of 28 participants.

Misery scores were low, with 26 of 28 participants reporting no symptoms (0 or 1 on the 11-point scale) and 2 of 28 participants reporting maximum scores of 3 and 4, respectively, at any time during the experiment. One of these two participants indicated feeling ill already before the experiment.

Figure 7 shows the percentage of trials in which participants held down the response button. Interestingly, a large portion of the participants (approximately 70%) did not hold down the response button at the beginning of the trial, even though the AV was still far away. For the non-yielding AV (Figure 7, left), participants released the response button as the AV came closer. However, about 13% of the participants did not have the response button released while the vehicle passed.

For the yielding AVs (Figure 7, right), participants gradually started pressing the response button after the AR interface appeared. For the baseline condition, this happened somewhat later, which can be explained by the fact that the AV only started to slow down at 3.9 s (thus, 0.8 s after the AR interfaces became visible), i.e., in the baseline condition, participants could not anticipate that the AV was going to stop until it actually started to decelerate. From the button press rates in the grey interval, it can be seen that the *Planes on vehicle* and *Pedestrian lights HUD* were slightly more effective than the *Virtual fence* and *Pedestrian lights HUD*, followed by the baseline condition.

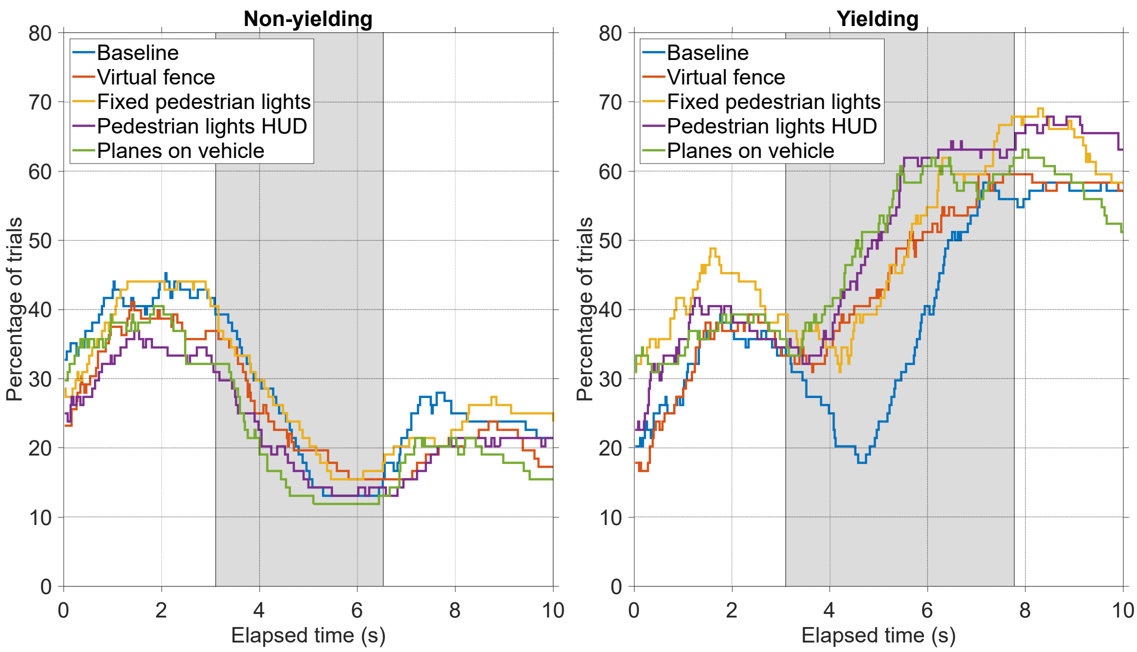

Figure 7. Button depress rates for the four AR interfaces and baseline condition, for non-yielding AVs (left), and yielding AVs (right). The grey background represents the interval between the moment the AR interface appeared (at 3.1 s) and the AV passed (non-yielding AV, 6.5 s) or the AV came to a full stop (yielding AV, 7.8 s).

Figure 8 shows the button press rates in the selected time interval for non-yielding AVs (a) and yielding AVs (b), as well as the mean composite scores calculated from 15 items of the post-experiment questionnaire (c), intuitiveness scores measured after each trial for non-yielding AVs (d) and yielding AVs (e), the preference rank (f), trust (g), timing (h), and size (i).

The results in Figure 8 show a fairly consistent picture in the sense that the baseline condition scored relatively poorly, as evidenced by a high button press rate for non-yielding AVs (a) and a low button press rate for yielding AVs (b), relatively low intuitiveness scores especially for yielding AVs (e), and on average the worst preference rank (f). Among the four AR interfaces, there were no major differences, although the *Fixed pedestrian lights* scored relatively poorly. This is evident from the poor button press rates (a), low composite score (c), low intuitiveness scores (d & e), and the poor preference rank (f) among the four AR interfaces.

The largest effects between the four interface conditions were, however, observed regarding the perception of the physical variables timing and especially size). More specifically the *Fixed traffic lights* were perceived as clearly too small (Figure 8i), and related to this, participants also believed that the traffic lights became visible too late (Figure 8h); its timing was identical to that of the other three interfaces, but this assessment might be due to the fact that it took participants extra time to visually locate it, creating the impression that it became visible too late.

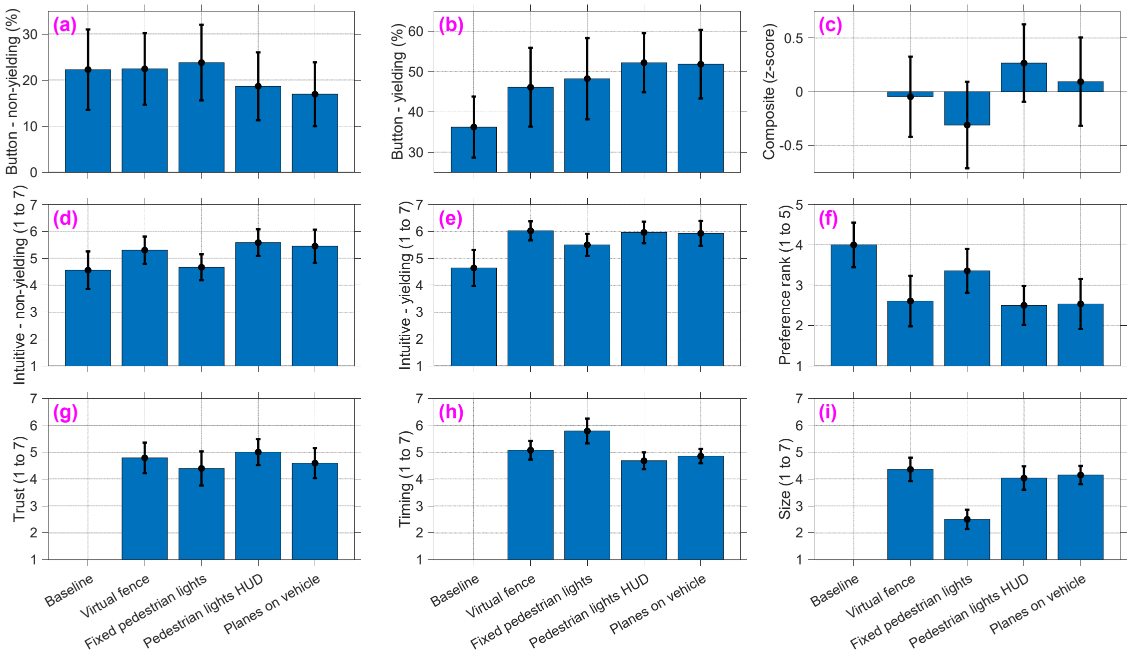

Figure 8. Means and 95% confidence intervals of participants' **(a)** button press rates for non-yielding AVs (%), **(b)** button press rates for yielding AVs (%), **(c)** composite scores based on post-experiment self-reports (*z*-score), **(d)** post-trial intuitiveness scores for non-yielding AVs (1: fully disagree, 7: fully agree), **(e)** post-trial intuitiveness scores for yielding AVs (1: fully disagree, 7: fully agree), **(f)** preference ranking (1: most preferred, 5: least preferred), **(g)** trust for decision-making (1: fully disagree, 7: fully agree), **(h)** interface trigger timing (1: too early, 7: too late), and **(i)** interface size (1: too small, 7: too large).

The corresponding results of the repeated-measures ANOVAs for the nine sub-figures are shown in Table 1. The effect sizes (partial η^2) were not particularly strong for the button presses (a & b), the composite score (c), or the trust score (g). However, they were fairly strong and statistically significant for the post-trial intuitiveness ratings (d & e), the preference rank (f), and as mentioned above , the timing (h) and size (i).

Table 1
Results of repeated-measures ANOVAs for the AR interface comparisons shown in Figure 8.

Dependent measure	**Measurement moment**	***df***	***F***	***p***	**partial η^2**
(a) Button press rate, non-yielding AVs	During trials	4,108	0.57	0.684	0.02
(b) Button press rate, yielding AVs	During trials	4,108	2.33	0.060	0.08
(c) Composite score	Post-experiment	3,78	1.65	0.185	0.06
(d) Intuitiveness, non-yielding AVs	Post-trial	4,108	2.88	0.026	0.10

(e)	Intuitiveness, yielding AVs	Post trial	4,108	6.34	< 0.001	0.19
(f)	Preference rank	Post-experiment	4,108	5.72	< 0.001	0.17
(g)	Trust	Post-experiment	3,78	0.91	0.441	0.03
(h)	Timing	Post-experiment	3,78	6.90	< 0.001	0.21
(i)	Size	Post-experiment	3,78	19.1	< 0.001	0.42

The results of the post-block interviews are indicated below:

- **Baseline.** Without an interface, many participants felt the situation was more akin to real-life scenarios. They indicated that the absence of clear signals or feelings of insecurity without the interface caused them to take longer in making decisions. The primary factors influencing their decision-making were the AV's deceleration and speed. Additionally, the changing pitch of the AV's sound during deceleration was reported to be a relevant cue.

  “Well, at some point you see that car slowing down. Then you still remain a bit apprehensive, thinking, ‘Okay, it's really stopping.’ And only then do I decide to start walking.”

- **Virtual fence.** Participants mentioned that the size of the Virtual Fence made it stand out, and some even found it somewhat intimidating. The interface was generally perceived as clear, particularly the green signal. However, the red signal was confusing for some participants; they were unsure whether it warned them not to cross or indicated that the AV would stop. The zebra crossing within the interface also posed a dilemma: it seemed to encourage walking, but this was in conflict with the red walls. Additionally, some participants expressed concerns that the interface might obstruct the view of other potential road users.

  “*Yes, so I found it a bit difficult at first, because you initially have those walls, which are red. So then you think, oh, is it red for them? And are you protected by those walls or something? And you're also in a zebra crossing. When I see a zebra crossing, I think, oh, I need to walk. So I found that a bit counterintuitive. But once you get used to it, I think you see it faster.*”

- **Fixed pedestrian lights.** The *Fixed pedestrian lights* were a familiar concept to participants. However, its appearance, timing, and placement posed challenges. Some participants reported they failed to notice it, especially during their first trial. Furthermore, its sudden appearance made it hard for participants to rely on. Additionally, a recurrent concern among participants was the need to switch gaze between the AV and the traffic light. Finally, participants suggested making the distinction between red and green clearer.

  “*More unclear, because the pole only appears relatively late. So you're really only focused on the car that's approaching. And at first, I either saw the pole or I didn't. You see the car approaching sooner, so then you start considering whether you're going to cross or not. And only then do you see the pole.*”

- **Pedestrian lights HUD.** The *Pedestrian lights HUD* also provided a familiar interface. Its upper-right positioning in their FOV was reported to be both an advantage and a disadvantage. Specifically, some participants reported that the HUD required them to roll their eyes, which felt unnatural to them, while others appreciated the constant visibility and fixed position within their FOV. Its sudden appearance sometimes led to initial startles or distractions, and some participants commented that it was obstructive and blocked a portion of their FOV. Finally, as with the *Fixed pedestrian lights*, according to some participants, the distinction between the red and green signals (lit-up vs. non-lit-up state) could have been clearer.

  “Very clear, indeed. I'm not used to it, so that's why the first one might have been a bit unclear. Because I was thinking, should I now look or should I clearly know what the car is going to do? But after that, it's very clear, because it's just close by and you actually see it right away.”

- **Planes on vehicle.** Participants initially found the *Planes on Vehicle* novel and took some time to become accustomed to it. The green signal and icon were perceived as clear. However, there was some uncertainty regarding the red signal, with participants unsure whether the AV would stop. Many felt that the *Planes on Vehicle* improved trust because the communication came directly from the AV. One colour blind person, however, did not immediately understand the meaning of this interface. Furthermore, some participants remarked that in hypothetical situations involving multiple vehicles, the effectiveness of the interface might decrease. It was suggested that, for optimal effectiveness in busier scenarios, all cars might need to adopt such an interface.

  "*I don't know, it made me a bit nervous, because I thought okay... It seems very illogical that there's a very large sign in front of the car.*"

  "Better than the previous one. I liked that you only have to pay attention to one thing. So you just have to look at the car and it's clear whether you can cross or not. Instead of having to look somewhere else for a signal. It was just clearly on the car."

Finally, a recurring topic regarding the timing of the AR interfaces was that they were activated somewhat late (even though this was still 0.8 s before the vehicle began to brake). Participants suggested that the AR interfaces could appear earlier, possibly in a default state, so that they would know where to look in advance to make their crossing decision. Another theme that emerged from the interviews is that participants tended to verify the advice of the AR interfaces with the movement of the vehicle.

Discussion

This study tested four AR interfaces for supporting the interaction with AVs among 28 human participants, complemented with a baseline condition without AR interface. For the research, we used actual AR in an outdoor environment, in contrast to previous research that used AR-in-VR. In our paradigm, participants stood outside and had to press a button as long as they felt safe to cross. The experiment was set up so that the AV was virtual; this way, we could offer the same vehicle movement and AR activation timings in every trial.

Some of the results did *not* match our previous studies. Specifically, a previous experiment in a CAVE environment (Tabone et al., 2023b) and a large-scale online survey (Tabone et al., 2023a) showed that among the AR interfaces tested, the world-locked *Fixed pedestrian lights* and head-locked *Pedestrian lights HUD* yielded relatively high intuitiveness ratings while the vehicle-locked *Planes on vehicle* received lower ratings, though still net positive. However, in the current study, it was found that the *Planes on vehicle* were rated highly, at around 6 on the the from 1 to 7, while the *Fixed pedestrian lights* interface was rated somewhat lower, at 4.7 and 5.5 for non-yielding and yielding AVs, respectively (see Figures 8d & 8e).

Although the relative rankings of the AR interfaces did not immediately correspond with prior research, the information-processing mechanisms did. For example, the observed phenomenon that a crosswalk in combination with the colour red causes confusion was also found in earlier research on the *Virtual fence* (Tabone & De Winter, 2023). The fact that a HUD and walls risk occlusion of other road users has previously been documented as well (Tabone & De Winter, 2023; Tabone et al., 2023b). Additionally, the fact that there can be an egocentric vs. exocentric perspective confusion when a car emits a red signal is also known (Bazilinskyy et al., 2020), and the problem that world-locked interfaces, like the fixed pedestrian lights in our case, cause divided attention, has also already been documented (e.g., Peereboom, Tabone et al., 2023). The explanation for the relatively poor performance of the *Fixed traffic lights* seems to lie in more practical factors: We had placed it relatively far away from the participant, at 15 m, and the red and green signals were not easy to distinguish from each other. This research thus shows that basic design decisions related to salience can have large effects on the effectiveness of AR interfaces.

The *Pedestrian lights HUD* also suffered from 'practical limitations'; it was positioned slightly off-centre, leading to some discomfort as individuals had to adjust their gaze (see Plabst et al., 2022, for a similar phenomenon). Note that participants could only see the HUD if they rotated their eyes, as the HUD always followed the user's head movement. Positioning the HUD towards a more central position would likely increase the risk of occlusion of relevant objects, which are typically centrally located in users' fields of view. In our study, participants experienced fairly low sickness scores, which differs from the higher scores observed in Peereboom et al.'s (2023) HUD study for pedestrians in VR; this difference is likely attributed to our HUD being placed at a larger distance (1.0 m, compared to their 0.36 m in Peereboom et al.), reducing the accommodation-vergence conflict. However, it should be anticipated that poorer task performance and/or nausea can occur in more dynamic situations, where the AR information is not locked to, or embedded in, the world. This concerns the use of HUDs while the user is walking through the environment, resulting in a moving background relative to static AR stimuli in the field of vision (Fukushima et al., 2020) or when standing still while the AR stimuli move in the field of view (Kaufeld et al., 2022). For future research, we suggest creating a HUD that does not require the user's direct focus. For instance, using a bright light, which can be noticed through peripheral vision, might offer a more comfortable experience for the user (e.g., Chaturvedi et al., 2019).

The *Fixed pedestrian lights* had the disadvantage that it required divided attention. This same disadvantage was mentioned in earlier experimental research in a CAVE environment, where participants took a relatively long time to visually identify the traffic lights (Tabone et al., 2023b). It might be the case that in the current real-world study, the task of identifying the traffic lights and the AV was even more challenging than in VR, because there is more visual clutter in the real world, such as other pedestrians and parked vehicles.

Indeed, we noticed that participants seemed to have difficulty with the experiment. For example, some participants had difficulty localising the attention-attractor circles or had to be reminded to look into these circles; participants also occasionally had to be reminded to press the response button if they felt safe to cross. One explanation is that we did not explicitly provide 'press now' instructions before each trial (e.g., Peereboom et al., 2023). Another explanation for the low depress rates is that the button, which operated through infrared light, might not have worked reliably in the outdoor environment[1]. Regardless, the response quality from the button presses is rather poor, with some participants not releasing the button when the AV passed (see Figure 7). The button-press results, as shown in Figure 7, resemble online crowdsourcing research on pedestrian-crossing decisions (e.g., Bazilinskyy et al., 2021, 2022; Sripada et al., 2021), where a portion of crowdworkers were apparently inattentive. One possible explanation for these results is that mental demands are higher with AR in the real world compared to AR-in-VR, because the real world is more cluttered, and the participant has more to keep track of, including perceiving real objects as well as maintaining postural stability and safety in the physical world. In turn, these observations suggest that simplicity and clarity of AR communication are likely even more important than in virtual or online experiments.

A limitation of our study is that it used simulated AVs; in future traffic, actual AVs would need to communicate their stopping intentions wirelessly to the pedestrian's AR headset. Furthermore, instead of outside-in tracking using base stations that emit infrared light, as used in the current study, there would likely need to be a form of inside-out tracking, where the pedestrian's headset detects objects in the environment (e.g., Bhakar et al., 2022). Furthermore, although the Varjo XR-3 is a state-of-the-art AR device, there were some technical hiccups, such as occasional loss of tracking, which caused VR objects to display an abrupt rotation, leading to some confusion or disorientation. Furthermore, while the headset offered a large FOV, it was still more limited than natural vision without a headset. Finally, in the current study, only a single AV came from the right. Using multiple vehicles would increase realism, something that is especially relevant for vehicle-locked AR interfaces. On the other hand, the environment was realistic, with university employees and students walking around occasionally, as well as some maintenance vehicles being present. Adding false positives, like a green-coloured signal in combination with an AV that does not stop, could be useful to investigate (over-)reliance on the AR interfaces (see also Holländer et al., 2019; Kaleefathullah et al., 2022).

Conclusion

In this study, different AR interfaces for AV-pedestrian interactions were assessed in an outdoor setting, distinguishing it from previous online and AR-in-VR research. The results showed that

[1] Excluding trials where the participant did not press the button at all in a given trial is not advisable as this could introduce bias because it is unknown if participants genuinely did not feel safe to cross or if there was a technical issue.

having an AR interface was generally preferred to no AR interface. Additionally, the results of post-trial interviews replicated previous information-processing-related findings, such as that a red-coloured surface can be confusing, since this cue can pertain to the AV or to the pedestrian. The experiment also showed that visual attention mechanisms are of key importance, with a world-locked traffic light causing challenges since pedestrians need to distribute attention between the approaching AV and the traffic light. Our findings also highlighted the importance of practical design considerations, such as placement and salience, in determining AR interface effectiveness. Finally, our results and observations suggest that participants found the present task challenging, which, given the complex and cluttered nature of the real world as opposed to virtual environments, points to the need for simplicity and clarity in AR-based communication.

Acknowledgements

This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 860410 (project name: SHAPE-IT).

References

Apple. (2023). Introducing Apple Vision Pro: Apple’s first spatial computer. https://www.apple.com/newsroom/2023/06/introducing-apple-vision-pro

Bazilinskyy, P., Dodou, D., & De Winter, J. C. F. (2020). External Human-Machine Interfaces: Which of 729 colors is best for signaling ‘Please (do not) cross’? *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Toronto, Canada, 3721–3728. https://doi.org/10.1109/SMC42975.2020.9282998

Bazilinskyy, P., Kooijman, L., Dodou, D., & De Winter, J. C. F. (2021). How should external Human-Machine Interfaces behave? Examining the effects of colour, position, message, activation distance, vehicle yielding, and visual distraction among 1,434 participants. *Applied Ergonomics*, *95*, 103450. https://doi.org/10.1016/j.apergo.2021.103450

Berge, S. H., Hagenzieker, M., Farah, H., & De Winter, J. (2022). Do cyclists need HMIs in future automated traffic? An interview study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *84*, 33–52. https://doi.org/10.1016/j.trf.2021.11.013

Bhakar, S., Bhatt, D. P., Dhaka, V. S., & Sarma, Y. K. (2022). A review on classifications of tracking systems in augmented reality. *Journal of Physics: Conference Series* (Vol. 2161, No. 1, p. 012077). IOP Publishing. http://doi.org/10.1088/1742-6596/2161/1/012077

Bindschädel, J., Krems, I., & Kiesel, A. (2022). Two-step communication for the interaction between automated vehicles and pedestrians. *Transportation Research Part F: Traffic Psychology and Behaviour*, *90*, 136–150. https://doi.org/10.1016/j.trf.2022.08.016

Bokc, T., Maurer, M., & Farber, G. (2007). Validation of the vehicle in the loop (vil); a milestone for the simulation of driver assistance systems. *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, 612–617. https://doi.org/10.1109/IVS.2007.4290183

Bos, J. E. (2015). Less sickness with more motion and/or mental distraction. *Journal of Vestibular Research: Equilibrium and Orientation*, *25*, 23–33. https://doi.org/10.3233/ves-150541

Butenuth, M., Kallweit, R., & Prescher, P. (2017). Vehicle-in-the-loop real-world vehicle tests combined with virtual scenarios. *ATZ Worldwide*, *119*, 52–55. https://doi.org/10.1007/s38311-017-0082-4

Chaturvedi, I., Bijarbooneh, F. H., Braud, T., & Hui, P. (2019). Peripheral vision: a new killer app for smart glasses. *Proceedings of the 24th International Conference on Intelligent User Interfaces*, Marina del Ray, CA, 625–636. https://doi.org/10.1145/3301275.3302263

Colley, M., Walch, M., & Rukzio, E. (2020). Unveiling the lack of scalability in research on external communication of autonomous vehicles. *Extended abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu, HI. https://doi.org/10.1145/3334480.3382865

Dey, D., Ackermans, S., Martens, M., Pfleging, B., & Terken, J. (2022). Interactions of automated vehicles with road users. In A. Riener, M. Jeon, & I. Alvarez (Eds.), *User experience design in the era of automated driving* (pp. 533–581). Cham: Springer. https://doi.org/10.1007/978-3-030-77726-5_20

Dey, D., Holländer, K., Berger, M., Eggen, B., Martens, M., Pfleging, B., & Terken, J. (2020). Distance-dependent eHMIs for the interaction between automated vehicles and pedestrians. *Proceedings of the 12th International Conference on Automotive User Interfaces and*

Interactive Vehicular Applications, Virtual Event DC, 192–204. https://doi.org/10.1145/3409120.3410642
Feng, Y., Yu, C., Xu, S., Liu, H. X., & Peng, H. (2018). An augmented reality environment for connected and automated vehicle testing and evaluation. *Proceedings of the 2018 IEEE Intelligent Vehicles Symposium*, Changshu, China, 1549–1554. https://doi.org/10.1109/IVS.2018.8500545
Final Form Studio. (2021). Sedan Car - 01. https://assetstore.unity.com/packages/3d/vehicles/land/sedan-car-01-190629
Fukushima, S., Hamada, T., & Hautasaari, A. (2020). Comparing world and screen coordinate systems in optical see-through head-mounted displays for text readability while walking. *Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality*, 649–658, Porto de Galinhas, Brazil. https://doi.org/10.1109/ISMAR50242.2020.00093
Gelbal, S. Y., Cantas, M. R., Guvenc, B. A., Guvenc, L., Surnilla, G., & Zhang, H. (2023). *Mobile safety application for pedestrians.* arXiv. https://doi.org/10.48550/arXiv.2305.17575
Google. (2020). A new sense of direction with Live View. https://blog.google/products/maps/new-sense-direction-live-view
Hasan, R., & Hasan, R. (2022). Pedestrian safety using the Internet of Things and sensors: Issues, challenges, and open problems. *Future Generation Computer Systems*, *134*, 187–203. https://doi.org/10.1016/j.future.2022.03.036
Hesenius, M., Börsting, I., Meyer, O., & Gruhn, V. (2018). Don't panic! Guiding pedestrians in autonomous traffic with augmented reality. *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, Barcelona, Spain, 261–268. https://doi.org/10.1145/3236112.3236148
Holländer, K., Wintersberger, P., & Butz, A. (2019). Overtrust in external cues of automated vehicles: An experimental investigation. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Utrecht, The Netherlands, 211–221. https://doi.org/10.1145/3342197.3344528
HTC VIVE. (2019). SteamVR Base Station 2.0. https://www.vive.com/eu/accessory/base-station2
Hussain, K. F., Radwan, E., & Moussa, G. S. (2013). Augmented reality experiment: drivers' behavior at an unsignalized intersection. *IEEE Transactions on Intelligent Transportation Systems*, *14*, 608–617. https://doi.org/10.1109/TITS.2012.2226239
Jungmann, A., Lang, C., Pinsker, F., Kallweit, R., Taubenreuther, M., & Butenuth, M. (2020). Artificial intelligence for automated driving – quo vadis? In T. Bertram (Ed.), *Automatisiertes Fahren 2019: Von der Fahrerassistenz zum autonomen Fahren 5. Internationale ATZ-Fachtagung* (pp. 117–134). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-27990-5_11
Kaleefathullah, A. A., Merat, N., Lee, Y. M., Eisma, Y. B., Madigan, R., Garcia, J., & De Winter, J. C. F. (2022). External Human-Machine Interfaces can be misleading: An examination of trust development and misuse in a CAVE-based pedestrian simulation environment. *Human Factors*, *64*, 1070–1085. https://doi.org/10.1177/0018720820970751
Kamalasanan, V., Mukbil, A., Sester, M., & Müller, J. P. (2022). Mixed reality agent-based framework for pedestrian-cyclist interaction. *Proceedings of the 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct*, Singapore, 363–368. https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00079
Kang, Y., Choi, S., An, E., Hwang, S., & Kim, S. (2023). Designing virtual agent human–machine interfaces depending on the communication and anthropomorphism levels in augmented reality. *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Ingolstadt, Germany, 191–201. https://doi.org/10.1145/3580585.3606460
Kaufeld, M., Mundt, M., Forst, S., & Hecht, H. (2022). Optical see-through augmented reality can induce severe motion sickness. Displays, 74, 102283. https://doi.org/10.1016/j.displa.2022.102283
Lakhdhir, S., Somanath, S., & Sharlin, E. (2023). Wearing awareness: Designing pedestrian-wearables for interactions with autonomous vehicles. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg, Germany. https://doi.org/10.1145/3544549.3585655
Lebeck, K., Ruth, K., Kohno, T., & Roesner, F. (2017). Securing augmented reality output. *Proceedings of the 2017 IEEE Symposium on Security and Privacy*, San Jose, CA, 320–337. https://doi.org/10.1109/SP.2017.13

Lingam, S. N., De Winter, J., Dong, Y., Tsapi, A., Van Arem, B., & Farah, H. (2023). *eHMI on the vehicle or on the infrastructure? A driving simulator study.* ResearchGate. https://www.researchgate.net/publication/362751474_eHMI_on_the_Vehicle_or_on_the_Infrastructure_A_Driving_Simulator_Study
Malik, J., Kim, N.-Y., Parr, M. D., Kearney, J. K., Plumert, J. M., & Rector, K. (2023). Do simulated augmented reality overlays influence street-crossing decisions for non-mobility-impaired older and younger adult pedestrians? *Human Factors*, 00187208231151280. https://doi.org/10.1177/00187208231151280
Maruhn, P., Dietrich, A., Prasch, L., & Schneider, S. (2020). Analyzing pedestrian behavior in augmented reality — Proof of concept. *Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Atlanta, GA. https://doi.org/10.1109/vr46266.2020.00051
Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64. https://doi.org/10.20982/tqmp.04.2.p061
Paiva, S., Ahad, M. A., Tripathi, G., Feroz, N., & Casalino, G. (2021). Enabling technologies for urban smart mobility: Recent trends, opportunities and challenges. *Sensors*, *21*, 2143. https://doi.org/10.3390/s21062143
Peereboom, J., Tabone, W., Dodou, D., & de Winter, J. (2023). *Head-locked, world-locked, or conformal diminished-reality? An examination of different AR solutions for pedestrian safety in occluded scenarios.* ResearchGate. https://www.researchgate.net/publication/371509780_Head-locked_world-locked_or_conformal_diminished-reality_An_examination_of_different_AR_solutions_for_pedestrian_safety_in_occluded_scenarios
Plabst, L., Oberdörfer, S., Ortega, F. R., & Niebling, F. (2022). Push the red button: Comparing notification placement with augmented and non-augmented tasks in AR. *Proceedings of the 2022 ACM Symposium on Spatial User Interaction*, Online, CA. https://doi.org/10.1145/3565970.3567701
Pratticò, F. G., Lamberti, F., Cannavò, A., Morra, L., & Montuschi, P. (2021). Comparing state-of-the-art and emerging augmented reality interfaces for autonomous vehicle-to-pedestrian communication. *IEEE Transactions on Vehicular Technology*, *70*, 1157–1168. https://doi.org/10.1109/TVT.2021.3054312
Qualtrics XM. (2023). Qualtrics. https://www.qualtrics.com
Sheridan, T. B. (2016). Recollections on presence beginnings, and some challenges for augmented and virtual reality. *Presence: Teleoperators and Virtual Environments*, *25*, 75–77. https://doi.org/10.1162/PRES_e_00247
Sripada, A., Bazilinskyy, P., & De Winter, J. C. F. (2021). Automated vehicles that communicate implicitly: Examining the use of lateral position within the lane. *Ergonomics*, *64*, 1416–1428. https://doi.org/10.1080/00140139.2021.1925353
Tabone, W., & De Winter, J. C. F. (2023). Using ChatGPT for human-computer interaction: A primer. *Royal Society Open Science*, *10*, 231053. https://doi.org/10.1098/rsos.231053
Tabone, W., De Winter, J. C. F., Ackermann, C., Bärgman, J., Baumann, M., Deb, S., Emmenegger, C., Habibovic, A., Hagenzieker, M., Hancock, P. A., Happee, R., Krems, J., Lee, J. D., Martens, M., Merat, N., Norman, D. A., Sheridan, T. B., & Stanton, N. A. (2021a). Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. *Transportation Research Interdisciplinary Perspectives*, *9*, 100293. https://doi.org/10.1016/j.trip.2020.100293
Tabone, W., Happee, R., Garcia, J., Lee, Y. M., Lupetti, M. C., Merat, N., & De Winter, J. C. F. (2023a). Augmented reality interfaces for pedestrian-vehicle interactions: An online study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *94*, 170–189. https://doi.org/10.1016/j.trf.2023.02.005
Tabone, W., Happee, R., Yang, Y., & De Winter, J. (2023b). *Immersive insights: Evaluating augmented reality interfaces for pedestrians in a CAVE-based experiment.* ResearchGate. https://www.researchgate.net/publication/370160064_Immersive_Insights_Evaluating_Augmented_Reality_Interfaces_for_Pedestrians_in_a_CAVE-Based_Experiment
Tabone, W., Lee, Y. M., Merat, N., Happee, R., & De Winter, J. C. F. (2021b). Towards future pedestrian-vehicle interactions: Introducing theoretically-supported AR prototypes. *13th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 209–218, Leeds, UK. https://doi.org/10.1145/3409118.3475149

Tong, Y., Jia, B., & Bao, S. (2021). An augmented warning system for pedestrians: user interface design and algorithm development. *Applied Sciences*, *11*, 7197. https://doi.org/10.3390/app11167197
Tran, T. T. M., Parker, C., & Tomitsch, M. (2023). Scoping out the scalability issues of autonomous vehicle-pedestrian interaction. *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Ingolstadt Germany, 167–177. https://doi.org/10.1145/3580585.3607167
Tran, T. T. M., Parker, C., Wang, Y., & Tomitsch, M. (2022). Designing wearable augmented reality concepts to support scalability in autonomous vehicle-pedestrian interaction. *Frontiers in Computer Science*, *4*. https://doi.org/10.3389/fcomp.2022.866516
Transparency Market Research. (2023). Automotive Head-up Display (HUD) market. https://www.transparencymarketresearch.com/automotive-head-up-display-market.html
Troel-Madec, M., Boissieux, L., Borkoswki, S., Vaufreydaz, D., Alaimo, J., Chatagnon, S., & Spalanzani, A. (2019). eHMI positioning for autonomous vehicle/pedestrians interaction. *Adjunct Proceedings of the 31st Conference on l'Interaction Homme-Machine*, Grenoble, France. https://doi.org/10.1145/3366551.3370340
United Nations. (2020). United Nations A/RES/74/299 (A/RES/74/299). Resolution adopted by the General Assembly on 31 August 2020. https://documents-dds-ny.un.org/doc/UNDOC/GEN/N20/226/30/PDF/N2022630.pdf?OpenElement
Unity. (2022). Unity 2021.3.13. https://unity.com/releases/editor/whats-new/2021.3.13
Van der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies*, *5*, 1–10. https://doi.org/10.1016/s0968-090x(96)00025-3
Varjo. (2023). Varjo XR-3 - The industry's highest resolution mixed reality headset. https://varjo.com/products/xr-3
Varjo Developer. (2023). Getting started with Varjo XR Plugin for Unity. https://developer.varjo.com/docs/unity-xr-sdk/getting-started-with-varjo-xr-plugin-for-unity
Varjo Technologies. (2023). Mixed reality – Varjo.com. https://varjo.com/use-center/get-to-know-your-headset/mixed-reality/#reference-marker
Wilbrink, M., Burgdorf, F. K. G., & Oehl, M. (2023). AR designs for eHMI–Communication between automated vehicles and pedestrians using augmented reality. Presented at *Humanist*, Braunschweig, Germany.
World Health Organization. (2018). *Global status report on road safety 2018: Summary.* https://apps.who.int/iris/handle/10665/277370
Zhang, X., Wu, X., Cools, R., Simeone, A. L., & Gruenefeld, U. (2023). ARcoustic: A mobile augmented reality system for seeing out-of-view traffic. *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Ingolstadt, Germany, 178–190. https://doi.org/10.1145/3580585.3606461

Supplementary Material

Figure S1. Attention-attractor circle at the left side of the participant.

Figure S2. Post-trial intuitiveness question.