Augmented reality interfaces for pedestrian-vehicle interactions

An online study

Tabone, Wilbert; Happee, Riender; García, Jorge; Lee, Yee Mun; Lupetti, Maria Luce; Merat, Natasha; de Winter, Joost

DOI
10.1016/j.trf.2023.02.005

Publication date
2023

Document Version
Final published version

Published in
Transportation Research Part F: Traffic Psychology and Behaviour

Citation (APA)
Tabone, W., Happee, R., García, J., Lee, Y. M., Lupetti, M. L., Merat, N., & de Winter, J. (2023). 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

Transportation Research Part F: Psychology and Behaviour 94 (2023) 170–189

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

Augmented reality interfaces for pedestrian-vehicle interactions: An online study

Wilbert Tabone [a,*], Riender Happee [a], Jorge García [b], Yee Mun Lee [b], Maria Luce Lupetti [c], Natasha Merat [b], Joost de Winter [a]

[a] Department of Cognitive Robotics, Faculty of Mechanical Engineering, Delft University of Technology, The Netherlands
[b] Institute for Transport Studies, Faculty of Environment, University of Leeds, The United Kingdom
[c] Department of Human Centred Design, Faculty of Industrial Design Engineering, Delft University of Technology, The Netherlands

ARTICLE INFO

Keywords:
Augmented reality
Pedestrian-vehicle interactions
Vulnerable road users
Automated vehicles
Online questionnaire
User study
Road crossing

ABSTRACT

Augmented Reality (AR) technology could be utilised to assist pedestrians in navigating safely through traffic. However, whether potential users would understand and use such AR solutions is currently unknown. Nine novel AR interfaces for pedestrian-vehicle communication, previously developed using an experience-based design method, were evaluated through an online questionnaire study completed by 992 respondents in Germany, the Netherlands, Norway, Sweden, and the United Kingdom. The AR indicated whether it was safe to cross the road in front of an approaching automated vehicle. Each interface was rated for its intuitiveness and convincingness, aesthetics, and usefulness. Moreover, comments were collected for qualitative analysis. The results indicated that interfaces that employed traditional design elements from existing traffic, and head-up displays, received the highest ratings overall. Statistical results also showed that there were no significant effects of country, age, and gender on interface acceptance. Thematic analysis of the textual comments offered detail on each interface design's stronger and weaker points, and revealed unintended effects of certain designs. In particular, some of the interfaces were commented on as being dangerous or scary, or were criticised that they could be misinterpreted in that they signal that something is wrong with the vehicle, or that they could occlude the view of the vehicle. The current findings highlight the limitations of experience-based design, and the importance of applying legacy design principles and involving target users in design and evaluation. Future research should be conducted in scenarios in which pedestrians actually interact with approaching vehicles.

1. Introduction

Future traffic, in which automated vehicles (AVs) will be driving in city environments, requires transparent communication of the intentions of the vehicle with interaction partners, such as vulnerable road users (VRUs). In traditional traffic, transparent communication between vehicles and vulnerable road users is achieved through implicit and explicit cues (Lee et al., 2021; Schieben et al., 2019). Implicit cues include vehicle speed, dynamics, and gap size, while explicit cues include the horn, hand gestures, and eye contact. VRUs base their crossing decisions primarily on implicit cues (Dey & Terken, 2017; Lee et al., 2021), whereas explicit cues

---

[*] Corresponding author.
E-mail address: w.tabone@tudelft.nl (W. Tabone).

https://doi.org/10.1016/j.trf.2023.02.005
Received 30 August 2022; Received in revised form 3 February 2023; Accepted 8 February 2023

tend to be used when implicit cues are ambiguous (Onkhar et al., 2021; Uttley et al., 2020). With the introduction of AVs in the urban environment, the lack of a driver or attentive passenger may require a different approach to communicating intent from the AV to the VRU (Ackermans et al., 2020; Carmona et al., 2021; Faas et al., 2020; Hensch et al., 2019). Several communication methodologies have been proposed to alleviate the problems of AV-VRU interactions. These include the use of smart road infrastructure (Löcken et al., 2019; Pompigna & Mauro, 2022; Toh et al., 2020), smart vehicle kinematics through the use of vehicle pitch, deceleration, and lateral position (Bindschädel et al., 2022; Dietrich et al., 2020; Fuest et al., 2018; Sripada et al., 2021), and external human–machine interfaces (eHMIs).

Various forms of eHMIs have been developed, including LED strips, LED screens, anthropomorphic elements, actuated robotic attachments, and projections on the road, amongst others (see Bazilinskyy et al., 2019; De Winter & Dodou, 2022; Dey, Habibovic, Pfleging, et al., 2020; Rouchitsas & Alm, 2019, for reviews of such interfaces). Despite their effectiveness in encouraging VRUs to (not) cross in front of the AV's path, current eHMI designs have some drawbacks, namely if the eHMI needs to signal to a single pedestrian in a group, or, for text-based eHMIs, if the message is in a language unfamiliar to the pedestrian. Furthermore, so far, there has been no standardisation of eHMIs, and therefore pedestrians may encounter a variety of different eHMIs on vehicles, which could cause confusion (Rasouli & Tsotsos, 2020; Tabone, De Winter, et al., 2021), with potentially dangerous consequences.

In an effort to address some of these problems, augmented reality (AR) has been proposed as a new type of communication in traffic. AR used by individual VRUs can alleviate several issues, especially the one-to-many communication problem, where multiple actors (vehicles and pedestrians) are present in the environment and it is not clear which actor is communicating to whom. Through

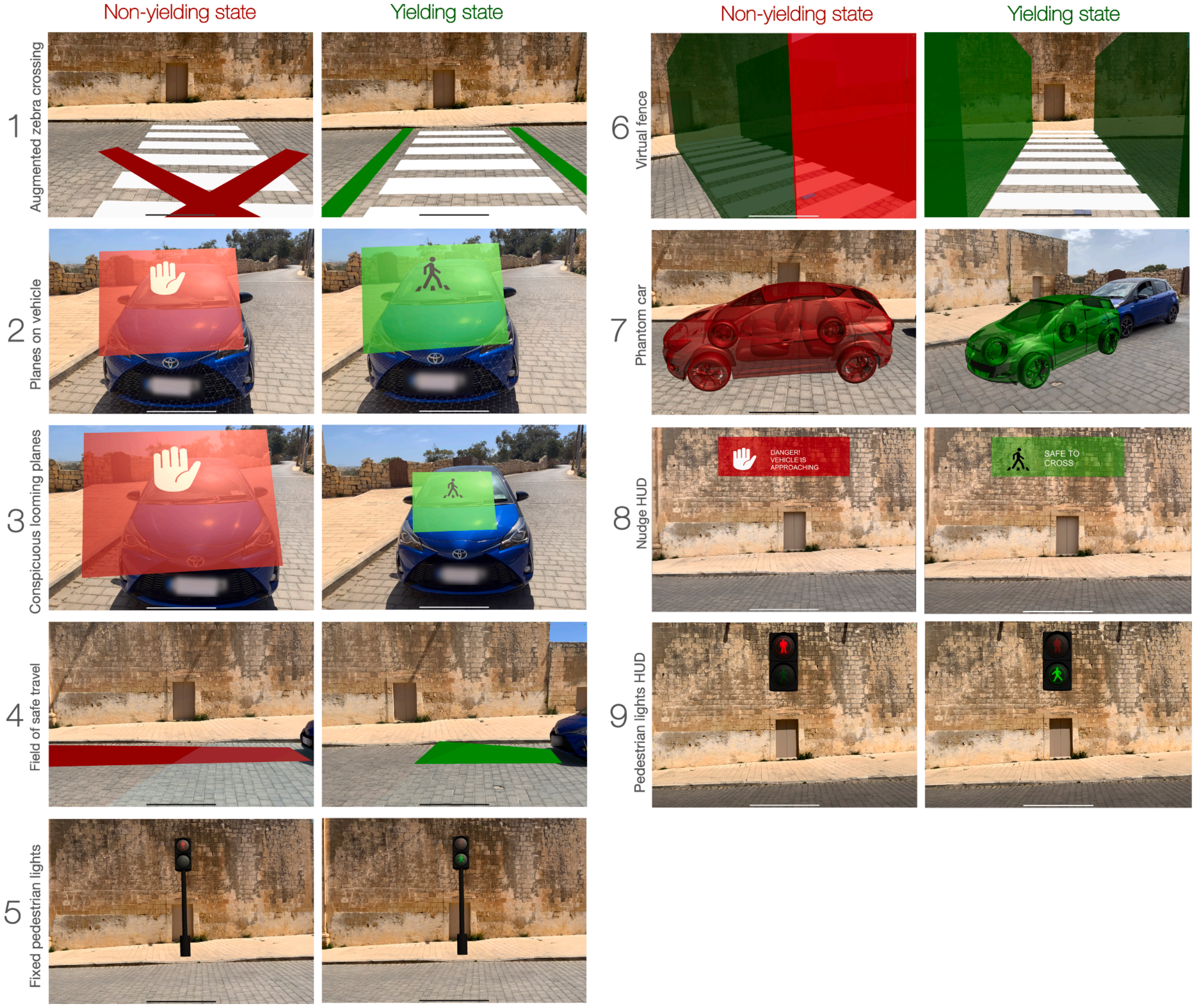


Fig. 1. The nine AR concepts for pedestrian-vehicle interactions designed and developed by Tabone, Lee, et al. (2021). In total, nine AR interface concepts were developed, each with a yielding and non-yielding state: 1. Augmented zebra crossing, 2. Planes on vehicle, 3. Conspicuous looming planes (i.e., planes which grew or shrank in size), 4. Field of safe travel, 5. Fixed pedestrian traffic lights, 6. Virtual fence, 7. Phantom car (i.e., a transparent car which indicates the vehicle's predicted future position), 8. Nudge HUD (i.e., a floating text message and icon which informed the pedestrian whether or not it was safe to cross), 9. Pedestrian traffic lights HUD. Interfaces 1, 4, 5, 6, and 7 are projected on the road surface, while Interfaces 2 and 3 are projected on the car. Interfaces 8 and 9 are head-locked, i.e., they remain in the user's field of view.

AR, the communication signal could be sent individually and separately to each pedestrian, and does not have to be constrained to the AV itself but can be presented anywhere in the environment (Tabone, Lee, et al., 2021; Tran et al., 2022).

So far, studies on AR for pedestrian-vehicle interaction consider the driver as the AR user, by highlighting pedestrians and/or cyclists in front of the vehicle (e.g., Calvi et al., 2020; Colley et al., 2021; Currano et al., 2021; Kim et al., 2018; Pichen et al., 2020). Such solutions are becoming technologically feasible when considering that the most recent vehicle models already feature AR-based head-up displays (Volkswagen, 2020). The use of AR by VRUs themselves is still relatively rare and has mostly been constrained to route navigation tasks (e.g., Bhorkar, 2017; Dancu et al., 2015; Dong et al., 2021; Ginters, 2019), for example as an add-on to Google Maps (Ranieri, 2020). Only a small, but growing number of studies have examined the use of AR for supporting VRUs in making safe crossing decisions. Examples include road projections such as zebra crossings, safe paths, and arrows (Hesenius et al., 2018; Li et al., 2022; Pratticò et al., 2021; Tran et al., 2022), visualisation of obstructed vehicles (Matviienko et al., 2022; Von Sawitzky et al., 2020), visualisation of collision times and conflict points (Tong & Jia, 2019), warning signs (Tong & Jia, 2019; Von Sawitzky et al., 2020), and car overlays (Tran et al., 2022). Using virtual reality, Oudshoorn et al. (2021) developed bioinspired eHMIs for pedestrian-AV interaction, whereas Mok et al. (2022) developed eHMIs in the form of laser-type rays emitted from the AV. The authors noted that these types of eHMIs may be hard to physically implement on real AVs, and that AR used by pedestrians (such as through AR glasses or handheld devices) could be a viable alternative.

It should be noted that most AR concepts for VRUs are still in a conceptual stage (videos, virtual reality), while only a few AR interfaces for VRUs have been demonstrated on a real road (Maruhn et al., 2020; Tabone, Lee, et al., 2021), or in a laboratory environment (Matviienko et al., 2022; Pratticò et al., 2021; Tran et al., 2022). In Tabone, Lee, et al. (2021), novel AR interfaces for pedestrian-AV interaction were developed and demonstrated in a real crossing environment. The interfaces were designed to assist pedestrians in the decision to cross the road in front of an approaching automated vehicle which was either yielding (stopping) or non-yielding. The interfaces were based on expert perspectives extracted from Tabone, De Winter, et al. (2021) and designed using theoretically-informed brainstorming sessions (see Fig. 1 for the interfaces). In total, nine AR interfaces were designed, each with a non-yielding and yielding state, depicted in red and green respectively. These colours were selected based on their high intuitiveness rating for signalling 'please (do not) cross' (Bazilinskyy et al., 2020).

Three of the interfaces were mapped to the road, four were mapped to the vehicle, and two were head-locked to the user's field of view. The ones mapped to the road were the *augmented zebra crossing*, which is a traditional zebra crossing design (1 in Fig. 1), *fixed pedestrian traffic lights* (5), which depicts a familiar pedestrian traffic light design across the road, and a *virtual fence* (6), which includes semi-translucent walls around a zebra-crossing and a gate that opens in the yielding state. The interfaces that were mapped to the vehicle included the *planes on the vehicle* (2), which displays a plane on the windshield area of the vehicle, the *conspicuous looming plane* (3), which grows or shrinks as the vehicle approaches the pedestrian depending on the AV's yielding state, the *field of safe travel* (4) which projects a field on the road in front of the vehicle to communicate safety, and the *phantom car* (7) which projects the vehicle's predicted future motion. The final two interfaces are head-up displays: the *nudge head-up display* (HUD) (8), which displays text and icons, and the *pedestrian lights HUD* (9), which displays a head-locked version of the pedestrian traffic lights.

In Tabone, Lee, et al. (2021), the interfaces were implemented on a handheld device (iPad Pro 2020) and demonstrated in a real crossing environment (Fig. 1), but no user study was performed. The concepts were designed using a 'genius'-based design approach (Saffer, 2010). In contrast to other design approaches, genius design does not involve users as part of the formal research phase. Instead, the design team relies on personal experience, existing knowledge of human behaviour, the problem space, and human cognition and psychology (Saffer, 2010). This approach offers the benefit of time efficiency, coherence of solutions with the original vision, and the flexibility to generate ideas quickly. Yet, such an approach could be contested as it addresses the problem space only from a designer's viewpoint without the involvement of the intended users (Nielsen, 2007).

Although a theoretical evaluation based on nine AR heuristics (Endsley et al., 2017) was performed in Tabone, Lee, et al. (2021), it is vital that AR concepts are evaluated empirically to assess whether the theoretically informed ideas are valid. Such an empirical evaluation would assess the viability of the 'genius' design approach in Tabone, Lee, et al. (2021) and whether the designers' intended effects would generalise to potential target users. Conducting a real-world study with the implemented AR prototypes would have been very difficult at the time of writing due to AR technology limitations, such as outdoor luminance levels that may hinder perception, latency issues that may lead to visually induced motion sickness, and ocular vergence-accommodation conflicts in open spaces (Buker et al., 2012; Rolland et al., 1995; Wann et al., 1995). Therefore, an online questionnaire study approach with a large number of participants was selected. A substantial number of previous works have conducted online user surveys to evaluate eHMIs for pedestrian-AV interaction (e.g., Bai, Legge, Young, Bao, & Zhou, 2021; Bazilinskyy, Dodou, & De Winter, 2020; Bazilinskyy, Kooijman, Dodou, & De Winter, 2021; Dey et al., 2020; Lau et al., 2021). However, no large-sample survey of AR interfaces for VRU-AV interactions has been conducted so far.

Hence, we attempt to fill this gap and build upon the previous design work reported in Tabone, Lee, et al (2021) by conducting an online video-based questionnaire study that investigates user acceptance of the AR interfaces across large numbers of participants, exploring key moderator variables (e.g., nationality, gender). Ratings of intuitiveness, convincingness, usefulness, aesthetics, and satisfaction with the interface were captured, which were thought to represent key dimensions of interface quality. These measures were based on previous studies which explored intuitiveness (Bazilinskyy et al., 2020), usefulness (Adell, 2010), quality of information (Lau et al., 2021), as well as aestheticism, attractiveness, and visibility (Métayer & Coeugnet, 2021). More specifically, it was reasoned that a high-quality AR interface should be easily understood (intuitive) and encourage people to follow up its recommendations (convincing), and be seen as useful in supporting pedestrian decision-making (usefulness). Furthermore, apart from encouraging performance, whether people like the AR interface (attractiveness, satisfaction) was seen as relevant, as when people might reject/disuse an (otherwise useful) AR interface on aesthetic grounds, it will still fail to be effective.

2. Method

In this study, participants viewed videos in a within-subject design, with 9 AR interfaces and 2 yielding behaviours. Participants rated each video according to a number of criteria. The video content, questionnaire design and procedures, and statistical analysis methods are explained below.

2.1. Videos

A total of 19 videos (at 30 fps) depicting an approaching AV with a representation of the AR interface in the virtual reality (VR) environment were created (Fig. 2). More specifically, nine videos depicted a yielding AV featuring a green-coloured (RGB: 32, 244, 0) AR interface, and nine videos depicted a non-yielding AV featuring a red-coloured (RGB: 244, 0, 0) AR interface.

A 19th video was created to depict a non-yielding AV without any interface. The latter was used at the start of the questionnaire to demonstrate how confusing and dangerous a situation without any form of signal would be, especially if the vehicle does not yield, while the other 18 videos were shown to participants in the experiment section of the questionnaire.

The videos were created based on a simulation created in a Unity-built VR environment (Unity, 2022). The road environment was obtained from previous research (e.g., Kaleefathullah et al., 2020) performed in the Highly Immersive Kinematic Experimental Research (HIKER) simulator located at the University of Leeds (University of Leeds, 2022). The videos mimicked the first-person view of a stationary pedestrian considering to cross in front of an approaching vehicle and looking to the right, on a one-way street. A one-way street was selected in order to standardise the direction of traffic flow, considering that the target population of the study were from countries with different traffic systems. Other studies focusing on road crossing have also utilised a one-way street scenario (e.g., Cavallo et al., 2019; Kaleefathullah et al., 2020; Weber et al., 2019).

Trigger points and speeds were adopted from a study on pedestrian crossing in the HIKER simulator (Kaleefathullah et al., 2020). The AV, represented by the same car model in each video, spawned out of sight from the field of view (Fig. 3, Point A) and moved at a constant speed of 30 mph (48 kph). All interfaces, irrespective of location and state, were triggered when the vehicle reached Point B, located 43 m from the participant (camera) location at Point E. For yielding AVs, the vehicle started decelerating at a rate of 2.99 m/s^2 at Point C, which is located 33 m from Point E, and it came to a full stop 3 m from Point E, at Point D. In the case of a non-yielding AV,

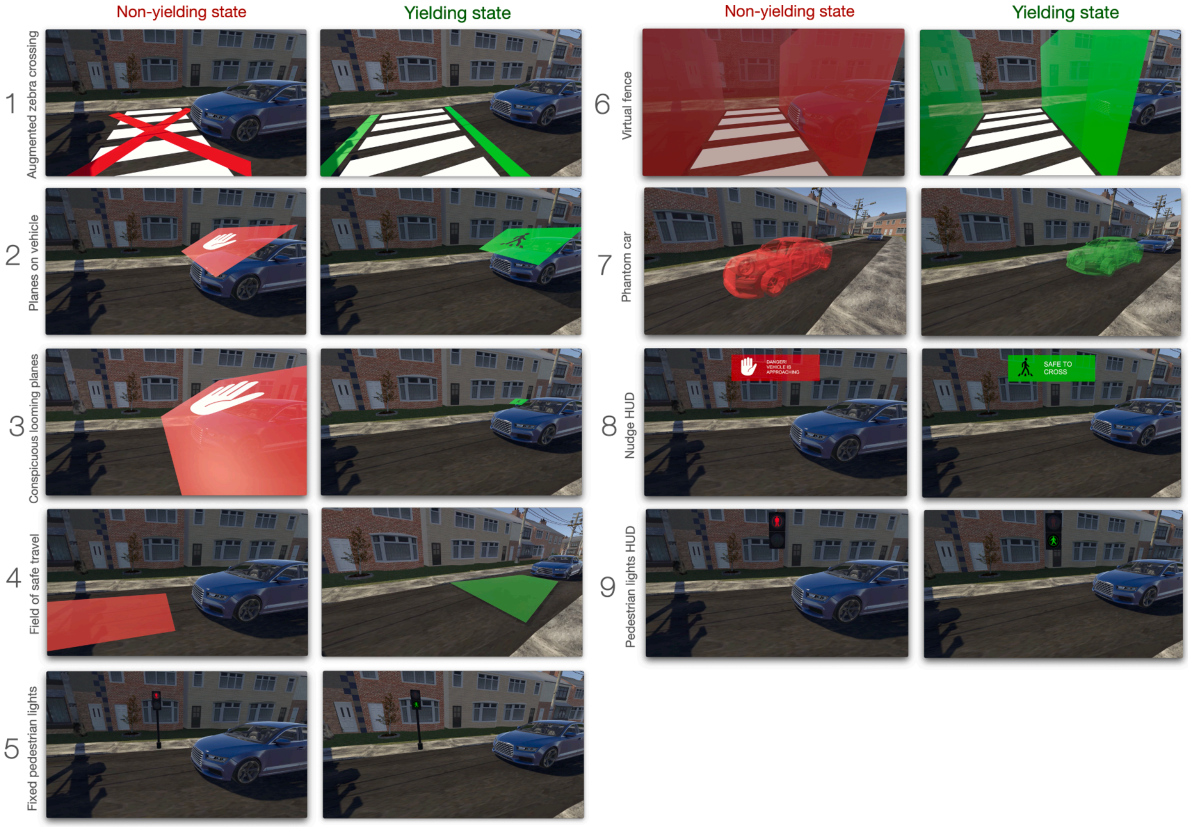


Fig. 2. The nine AR interfaces presented in a VR environment used for this online questionnaire study. Interfaces 1, 4, 5, 6, and 7 are projected on the road surface, while Interfaces 2 and 3 are projected on the car. Interfaces 8 and 9 are head-locked. The interfaces were adapted from Tabone, Lee, et al. (2021).

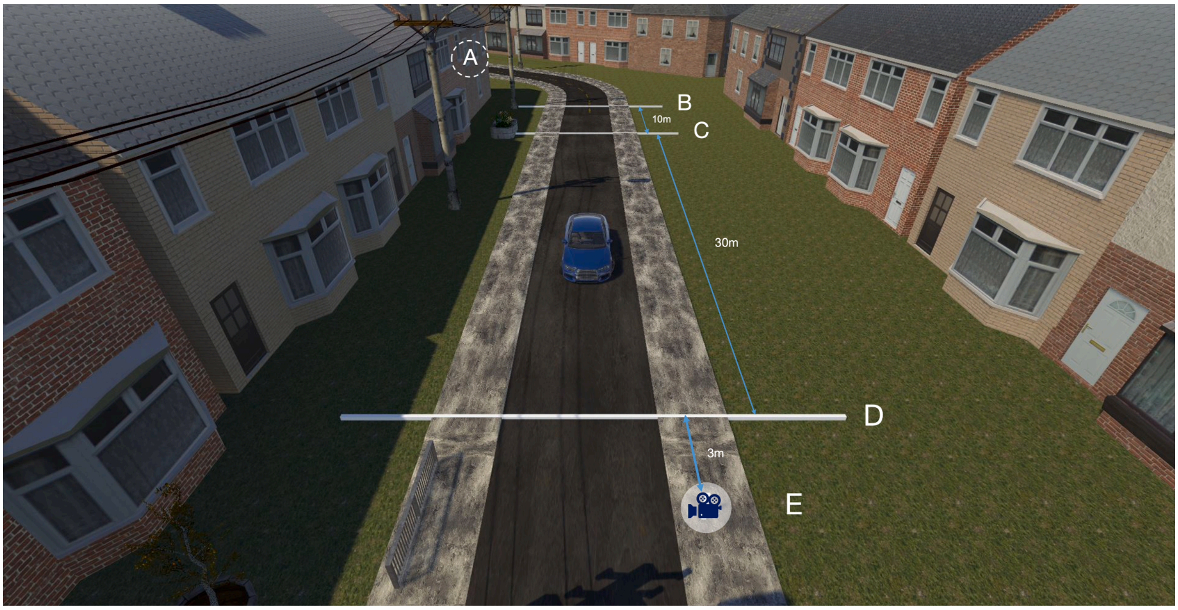


Fig. 3. Virtual environment used in the videos. Each salient point is demarcated by a label, together with the distance (in metres) between each point. A: spawn point, B: AR interface onset, C: AV deceleration onset, D: stopping point, E: participant location. The participant position is also marked with a camera icon.

the vehicle maintained its initial speed of 30 mph throughout.

Each video started with the camera pointing towards the other end of the crossing (Fig. 4, at time 0 s). The camera then slowly panned to the right as the vehicle approached from point A, starting at an elapsed time of 0.5 s. At an elapsed time of 2 s, the camera would have rotated by an angle of 45°, and the approaching vehicle and AR interface (regardless of type) could be seen simultaneously. At 4 s, the camera started to rotate back to the front-facing position, and it stopped rotating at 20° to the right for the yielding state (elapsed time: 9 s), and fully facing the front for the non-yielding AV (elapsed time: 8 s) so that the vehicle could be observed driving over the crossing area.

In addition to videos, side-by-side images were created per AR interface, for insertion in the questionnaire (see Fig. 5 for an example). For the yielding AV, the frame where the vehicle came to a complete stop was selected, while for the non-yielding state, the frame at an elapsed time of 6 s was used so that each screenshot had a similar perspective on the road. The only exception was the side-by-side comparison of the *phantom car*, where the screenshots were taken with respect to the location of the phantom car interface, rather than the actual vehicle, so that both the interface and the vehicle could be seen in the screenshots. The 19 videos produced for the experiment are included in the Supplementary Material.

2.2. Questionnaire procedure

The online questionnaire was administered to 1500 respondents from Germany, the Netherlands, Norway, Sweden, and the United Kingdom. These countries were selected based on the geographical locations of the participating partners of the Horizon 2020 SHAPE-

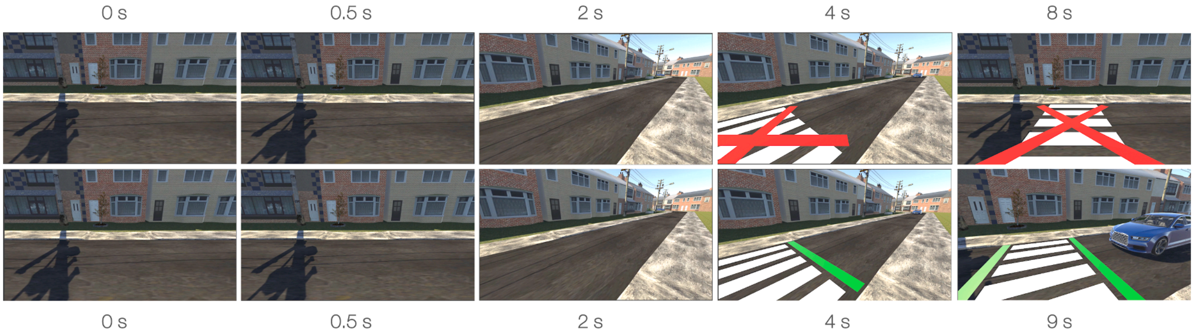


Fig. 4. Screenshot of the camera view for Augmented Zebra Crossing at key timestamps. The screenshot at the top are for the non-yielding state, while the bottom screenshots correspond to the yielding state.

Fig. 5. Example of side-by-side image for AR concept 1, Augmented zebra crossing. Left: non-yielding state, Right: yielding state.

IT project, which funded this research. These five European countries also have a strong research base in automated vehicle development (Hagenzieker et al., 2020) and are likely candidates for the early deployment of eHMIs and AR interfaces. The questionnaire was developed in English using the Qualtrics XM (Qualtrics, 2022) survey platform and distributed to representative Internet panels through the German market research institute INNOFACT AG (Innofact, 2022), which has been used in previous research on the acceptance of AVs (Nordhoff et al., 2021).

A screening questionnaire, prepared in the national language of each of the target respondents' countries, was added by INNOFACT, to control for age, gender, and nationality and filter out respondents who were uncomfortable with completing the questionnaire in English. Our requested target sample was an equal distribution across countries, gender, and split between five (18–29, 30–39, 40–49, 50–59, 60–69) age groups. INNOFACT ensured that participants only participated using a desktop device, and safeguards against bots and duplicate respondents were also taken.

The survey ran from February to April 2022, and the respondents were financially compensated with approximately €3. The study was approved by the Human Research Ethics Committee of the TU Delft under application number 1984.

2.3. Questionnaire design

2.3.1. Introductory information

First, a brief overview of AR and VR technologies was presented, together with examples of popular AR apps, so that the unfamiliar respondents would have a clearer picture of what would be discussed in the rest of the questionnaire. This was followed by an example of what the future could look like with the introduction of AR glasses, a brief introduction to the future urban environment, and the need for communication between AVs and pedestrians. The problem of having no clear signals from the car due to the lack of a driver was demonstrated through the baseline video (i.e., without AR interface) of a non-yielding AV. The respondents were provided with an explanation of the purpose of the study, where the potential of solving the communication issue using AR interfaces would be explored.

2.3.2. Consent

Respondents were provided with a consent section, which contained the experimenters' names, contacts, conditions to participate (being 18 years or older), the main purpose of the study, and the approximate length of the questionnaire (30 min). It was also highlighted that there were no risks associated with participation and that the questionnaire was anonymous and voluntary. Respondents were encouraged to close the page if they disagreed. Moreover, a question asking whether the instructions were read and understood was provided (Q1). If 'No' was selected, the questionnaire was terminated.

2.3.3. Demographics

Next, respondents were asked about their identifying gender (Q2), age (Q3), country of residence (Q4), and their highest level of formal education completed (Q5). Respondents were presented with the Affinity for Technology Interaction (ATI) scale (Franke et al., 2019) to gauge their affinity with technological systems (Q6). Respondents were then asked if they had ever used VR headsets (Q7) and AR apps (Q8), and how willing they would be to use AR wearables in general (Q9), specifically on the road as a pedestrian (Q10), and for the specific task of assisting pedestrians in crossing a road in front of an AV (Q11).

The respondents were then asked whether they had ever encountered AVs before (Q12), their daily walking time as pedestrians (Q13) (as used in Deb et al., 2017), and their primary mode of transportation (Q14). The last part in the demographic section treated any constraints in personal mobility (Q15) and included a colour blindness test (Q16) (Ishihara, 1917; as used in Bazilinskyy et al., 2020).

2.3.4. Video presentation of AR interfaces and rating questions

Following a brief introduction to the experiment, participants proceeded to the main part of the study, where the yielding and non-yielding state of the nine interfaces were presented, together with various rating questions.

The videos from each interface were presented on a separate page, having the title of the respective interface (see Fig. 2). The order in which the nine interfaces were presented was randomised for each respondent. Each interface page first presented the non-yielding-state video, followed by the yielding-state video. The videos auto-played and looped. All 18 videos were presented to each participant.

Below each video depicting a non-yielding AV, the respondents used a 7-point Likert scale (Strongly disagree, Disagree, Somewhat disagree, Neither agree nor disagree, Somewhat agree, agree, Strongly agree) to rate whether:

- "The interface in the video above is intuitive for signalling 'Please do NOT cross the road'" (Intuitiveness: Q17).
- "The interface in the video above convinced me NOT to cross the road" (Convincingness: Q18).

and below each video depicting a yielding AV, the following two questions were asked:

- "The interface in the video above is intuitive for signalling 'Please cross the road'" (Intuitiveness: Q19).
- "The interface in the video above convinced me to cross the road" (Convincingness: Q20).

Intuitiveness and convincingness were regarded as two key elements of interface quality, where the former refers to whether the message is readily understandable, and the latter refers to whether the interface would empower people to cross or not cross the road.

The video subsection containing the yielding and non-yielding videos along with the respective intuitiveness and convincingness items was followed by a side-by-side screenshot of the interface's states. A matrix table was presented with a 5-point descriptor scale (Q21) for interpretability, where the respondents had to rate the following:

- "Do you think that the interface was triggered too early or too late?" (too early – too late) (Q21.1)
- "Do you think that the interface is too small or too large? (too small – too large) (Q21.2)
- "How clear (understandable) was the interface to you?" (very unclear – very clear) (Q21.3)
- "How visually attractive is this interface to you?" (very unattractive – very attractive) (Q21.4)

Q17–Q21 were inspired from previous work which looked at perceived quality/clarity of information (Bazilinskyy, Dodou, & De Winter, 2020; Rahman, Lesch, Horrey, & Strawderman, 2017; Adell, 2010; Lau et al., 2021), and attractiveness, aestheticism, ease of understanding, and the adequacy of information, amongst others (Métayer & Coeugnet, 2021).

Each interface page ended with a 9-item acceptance scale (Van Der Laan et al., 1997) to collect further ratings on facets of usefulness and satisfaction (Q22.1–Q22.9). A free text area (Q23) was added to let respondents elaborate on their ratings, "Please add a few words to justify your choices above (eg. comment on the shape, colour, functionality, and the clarity of the interface)."

2.3.5. Final questions

The final section of the questionnaire opened with a question on whether such AR interfaces would be useful for crossing the road in future traffic (Q24). This query was followed by three side-by-side screenshots contrasting various interface elements, and the following three statements:

- "I prefer interfaces mapped to the street rather than on the vehicle" (Q25)
- "I prefer interfaces with text rather than interfaces with just graphical elements" (Q26)
- "I prefer interfaces that move around with my head rather than interfaces that stay fixed" (Q27), to which the respondent was answered with a 5-point Likert agreement scale from Strongly disagree to Strongly agree.

The penultimate question related to whether the respondent would like to have the ability to customise the interfaces (Q28). The final question once again asked whether the respondent would be willing to use such interfaces as an aid for crossing after having seen all examples, assuming that they own AR glasses (Q29).

2.4. Analysis

Mean item scores for the AR interfaces in their yielding and non-yielding states were computed and visualized in scatter plots, together with 95 % confidence intervals. The confidence intervals were computed by applying a correction for within-subjects effects of the nine AR interfaces, according to a method presented by Morey (2008).

Differences between ratings of AR interfaces were examined using a repeated-measures ANOVA with an alpha level of 0.05. This was followed by paired-samples *t*-tests. Here, an alpha value of 0.005 was used to reduce the occurrence of false positives, compared to the more commonly used alpha value of 0.05 (Benjamin et al., 2018). It should be noted that because our sample size was large, even small within-subject differences between the AR interfaces were strongly significant.

For the assessment of the effects of the moderator variables (gender, age group, educational level), a repeated-measures ANOVA was used with the AR interface as a within-subject variable and the moderator variable subgroup (e.g., male, female) as a between-subjects variable (alpha = 0.05). Additionally, statistical comparisons between ratings for AR interfaces between participant groups (e.g., males vs females) were performed using independent-samples *t*-tests (alpha = 0.005).

Apart from testing differences between AR interfaces and the effects of moderator variables, Pearson product-moment correlation coefficients among item scores were computed to evaluate redundancy among items. Highly correlated items were aggregated to form a composite score.

The textual responses were evaluated through thematic analysis (Kiger & Varpio, 2020). All text responses were read, with responses copied into a separate document if a common theme emerged. For example, if multiple participants commented that a particular interface was 'slow', then all comments with such a statement were extracted and placed in a text document under the section pertaining to the AR interface. Following the collation of all comments, four comments per interface (two per positive and two per negative valence were selected), depending on which theme was featured the most in that interface's comment section.

3. Results

In total, 1500 participants answered the questionnaire. An initial quality filtering process was carried out to remove respondents who did not complete the entire questionnaire ($n = 357$) or answered 'no' to the consent item (Q1) ($n = 39$). Next, the recorded duration in seconds was used to omit the top 10 % of fastest respondents (i.e., those who completed the questionnaire in 593 s or less, $n = 110$), since the fastest respondents are likely to yield relatively low-quality data (De Winter & Hancock, 2015). The resulting sample size was 992 (492 males, 491 females, 8 non-binary, 1 not specified). Within the resulting sample, the median time to complete the questionnaire was 23.3 min (25th percentile = 16.4 min, 75th percentile = 33.6 min).

General characteristics of the 992 retained respondents were as follows:

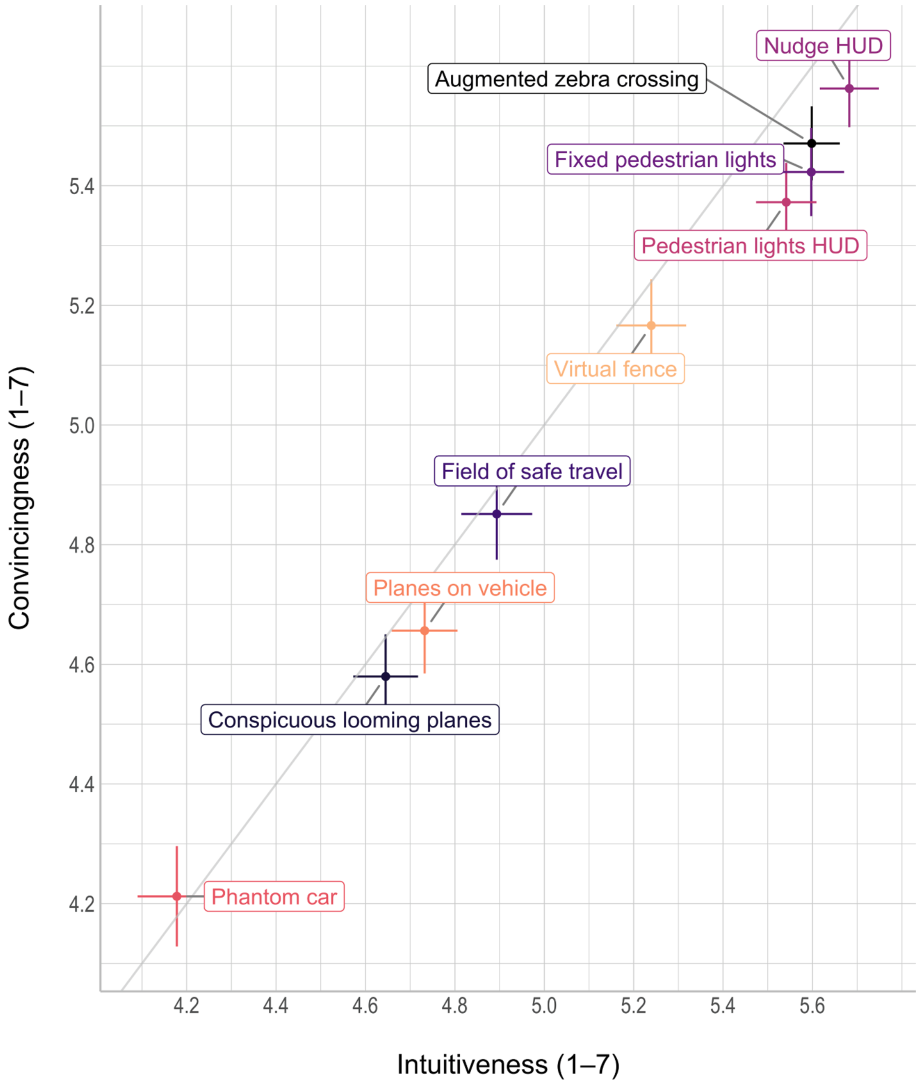


Fig. 6. Scatter plot of intuitiveness ratings (mean of Q17 and Q19) and convincingness ratings (mean of Q18 and Q20) per AR interface. In this figure, ratings for the yielding and non-yielding states were averaged. The error bars represent 95% confidence intervals.

- Country: 202 were from Germany, 197 were from the Netherlands, 184 were from Norway, 197 were from Sweden, and 212 were from the United Kingdom (Q4).
- Age: The age (Q3) ranged from 18 to 69 ($M = 45.10$, $SD = 14.17$).
- Education: 54 % ($n = 536$) indicated that they went to university, 25 % ($n = 246$) attended trade or vocational school, whereas 21 % ($n = 210$) indicated 'none of these' (Q5).
- Constraints: 17 % ($n = 170$) reported some form of mobility constraint (Q15).
- Constraints: 3 % ($n = 32$) were considered colour blind as they submitted three or more incorrect answers (Bazilinskyy et al., 2020) for the six-item Ishihara colour blindness test (Q16).

The results regarding AR and VR use indicated the following:

- 42 % of respondents had used a VR headset before (Q7).
- 45 % had used AR apps before (Q8).
- On a scale of 1 (Strongly unwilling) to 5 (Strongly willing), the mean response to "How willing would you be to use AR glasses?" (Q9) was 3.59 ($SD = 1.04$).
- For "How willing would you be to use AR glasses on the road as a pedestrian" (Q10), the mean was 3.10 ($SD = 1.13$).
- For "How willing would you be to use AR glasses on the road if these warn you about how safe it is to cross in front of a self-driving car?" (Q11), the mean was 3.30 ($SD = 1.12$).

Since the goal of this research was to perform a population-level evaluation of the AR interfaces, colour blind participants or participants with a mobility constraint were not excluded from the analysis.

3.1. *Ratings of videos depicting AR interfaces*

Table S1 in the Supplementary material shows the means across the 992 respondents for the 17 items for each of the nine AR interfaces. From this table, it can be seen that there are clear redundancies among the items, with some AR interfaces producing considerably higher ratings than others on almost all of the 17 items.

In an attempt to better understand item redundancy, several correlational analyses were performed. In particular, Fig. 6 shows the mean intuitiveness ratings (Q17, Q19) and convincingness ratings (Q18, Q20) for the nine AR interfaces. The ratings were very highly correlated ($r = 0.998$), indicating that the intuitiveness and convincingness questions yielded nearly the same information. Fig. 6 also shows that the *Nudge HUD* scored highest, followed by the *Augmented zebra crossing, Fixed pedestrian lights, Pedestrian lights HUD,* and *Virtual fence.* The *Phantom car* yielded the lowest ratings.

In the same vein, Fig. 7 shows the averaged intuitiveness and convincingness rating for the nine AR interfaces for yielding AVs versus non-yielding AVs. Again, a strong association ($r = 0.93$) is seen, indicating that the AR interfaces were rated similarly regardless of whether the vehicle was stopping or not. We performed a two-way repeated-measures ANOVA of the averaged intuitiveness and convincingness rating with AR interface and yielding state as within-subject factors. Results showed a significant effect of the AR interface, $F(8,7928) = 197.4, p < 0.001$, partial $\eta^2 = 0.17$, but not of yielding state $F(1, 991) = 0.12, p = 0.728$, partial $\eta^2 = 0.00$. There was, however, a significant AR interface × yielding state interaction, $F(8, 7928) = 41.5, p < 0.001$, partial $\eta^2 = 0.04$. Follow-up paired-samples *t*-tests showed that several AR interfaces (i.e., *Augmented zebra crossing, Field of safe travel, Fixed pedestrian lights, Nudge HUD, Pedestrian lights HUD*) yielded somewhat higher ratings for the non-yielding state than for the yielding state ($p < 0.005$ according to paired-samples *t*-tests). The *Virtual fence* and the *Planes on vehicle*, on the other hand, were rated statistically significantly higher for yielding AVs than for non-yielding AV.

A correlation matrix (Fig. 8) of the mean ratings for each interface revealed strong associations between all 17 measured items, except for the small/large item (Q21, Item 1) and early/late item (Q21, Item 2). The correlation coefficients between the means of the 15 other items ranged from $r = 0.862$ (for irritating/likeable [Q22, Item 6] vs sleep-inducing/raising alertness [Q22, Item 9]) to $r = 0.999$ (unpleasant/pleasant [Q22, Item 2] vs irritating/likeable [Q22, Item 6]).

3.2. *Descriptor scale (Q21), acceptance scale (Q22), and composite score*

Because correlation coefficients between items were very high, it was decided to compute a composite score of the 15 strongly-correlated items (unit-weight method, see DiStefano et al., 2009).[1] More specifically, for each AR interface, a 992 participant × 15 matrix was available. The matrices were concatenated, yielding an 8928 × 15 matrix, and subsequently standardised, so that the item mean was 0 and the standard deviation was 1. The scores of the 15 items were summed, thus producing an 8928-long vector, which was then standardised. Finally, the 8928-long vector was partitioned back to the nine interfaces, so that a composite score was available for each participant and AR interface. Fig. 8 shows that the mean composite score correlated very strongly with each of its

---

[1] An inspection of the eigenvalues of the correlation matrix (9 AR interfaces × 15 items) showed strong uni-dimensionality. More specifically, the first component explained 96.5% of the variance in the participant means, and the corresponding Cronbach's alpha value was 0.990. Additionally, the correlation matrix at the participant level (992 participants x 15 items) showed strong uni-dimensionality as well, with the first component explaining 67.6% of the variance in the means of the 9 AR interfaces, and Cronbach's alpha being 0.962.

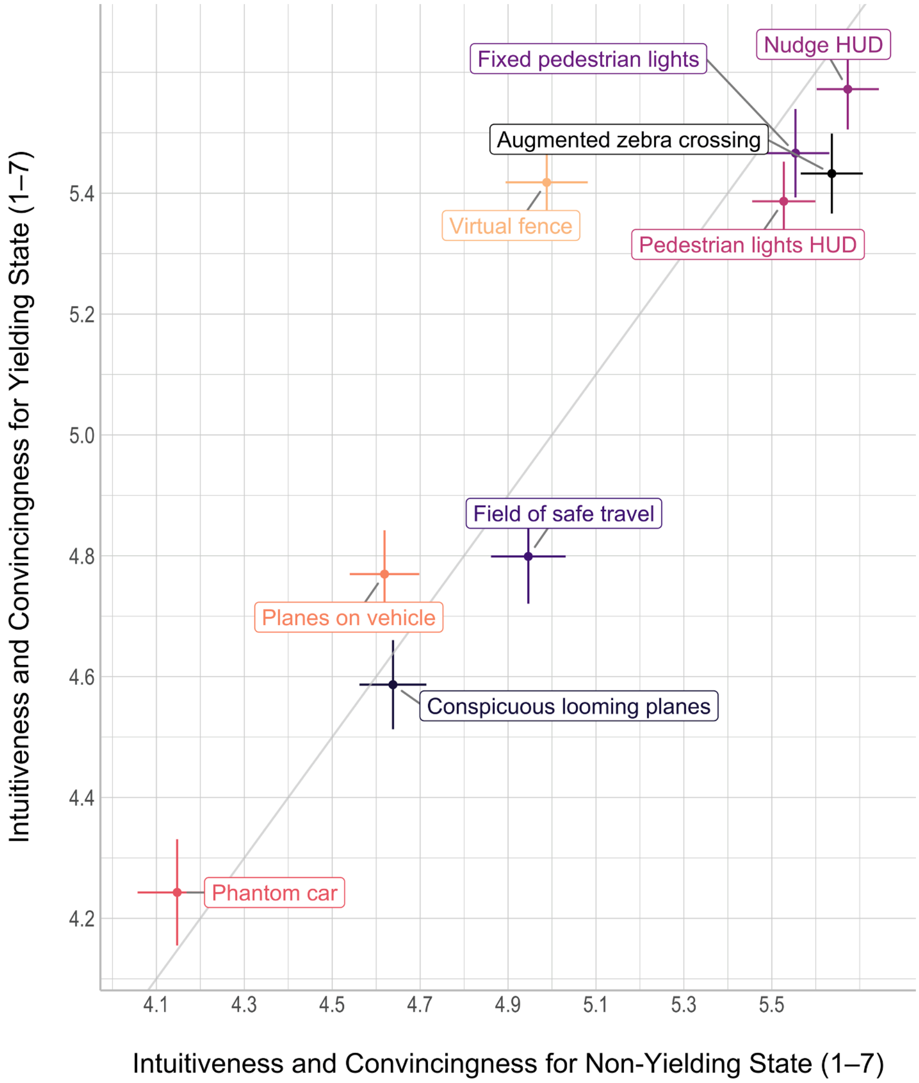


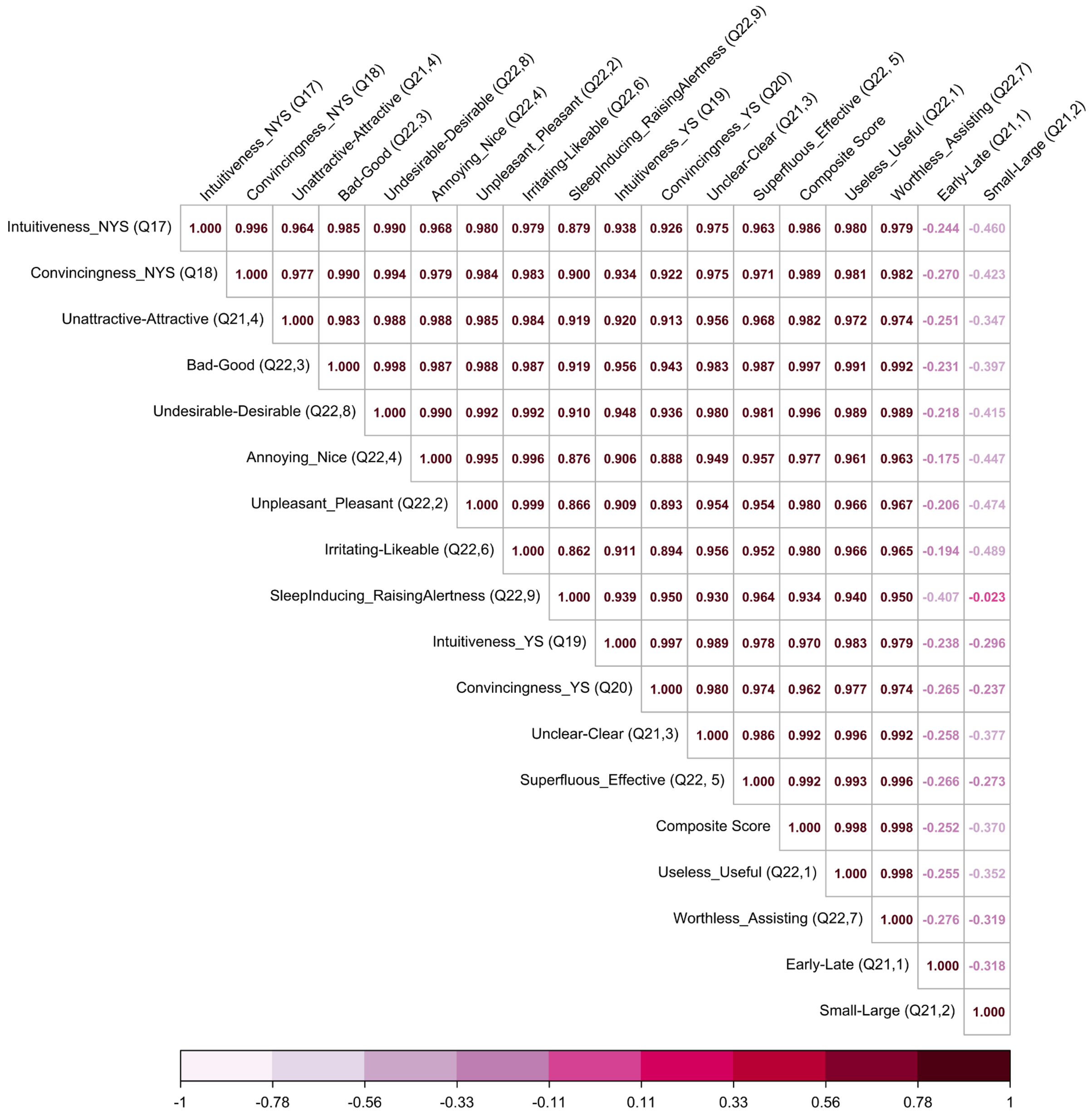
Fig. 7. Scatter plot of averaged intuitiveness and convincingness ratings of the yielding state (mean of Q19 & Q20) versus the non-yielding state (mean of Q17 & Q18) of each AR interface. The error bars represent 95% confidence intervals.

defining items, which confirms that the composite score captures a large amount of the variance (96.5 %) in the mean ratings of the nine AR interfaces. The strongest correlations between the composite score and the individual items ($r = 0.997, 0.998$) occurred for the items useful/useless (Q22.1), bad/good (Q22.3), and worthless/assisting (Q22.7). This suggests that the meaning of the composite score is well described by the colloquial phrase 'whether the AR interface is good or not'.

The mean and standard deviation of the composite score per AR interface are shown in Table 1. The findings align with the above results (Figs. 6 and 7) that the *Nudge HUD* was most favoured while the *Phantom car* was least favoured. A one-way repeated-measures ANOVA of the composite score showed a significant effect of the AR interface, $F(8,7928) = 195.0, p < 0.001$, partial $\eta^2 = 0.16$. A total of 32 of 36 pairs of AR interfaces were statistically significantly different from each other ($p < 0.005$), see Table 1.

3.3. Assessment of moderator variables

Gender: Fig. S2 in the supplementary material shows a strong correlation ($r = 0.980$) between the mean composite scores for male and female respondents. A repeated-measures ANOVA of the composite score, with the AR interface as a within-subject factor and gender (male or female) as a between-subjects factor showed a significant effect of AR interface, $F(8, 7848) = 192.6, p < 0.001$, partial $\eta^2 = 0.16$, and no significant effect of gender, $F(1, 981) = 0.36, p = 0.547$, partial $\eta^2 = 0.00$, but a significant AR interface × gender interaction, $F(8, 7848) = 2.00, p = 0.043$, partial $\eta^2 = 0.00$. The interaction effect was extremely small, however, and scores for the nine AR interfaces did not differ significantly between males and females. More specifically, independent-samples t-tests for the nine AR interfaces yielded p-values between 0.087 and 0.952 (*Conspicuous looming planes*: Mean (*SD*) males/females: −0.40 (1.03)/−0.29



	Intuitiveness_NYS (Q17)	Convincingness_NYS (Q18)	Unattractive-Attractive (Q21,4)	Bad-Good (Q22,3)	Undesirable-Desirable (Q22,8)	Annoying_Nice (Q22,4)	Unpleasant_Pleasant (Q22,2)	Irritating-Likeable (Q22,6)	SleepInducing_RaisingAlertness (Q22,9)	Intuitiveness_YS (Q19)	Convincingness_YS (Q20)	Unclear-Clear (Q21,3)	Superfluous_Effective (Q22, 5)	Composite Score	Useless_Useful (Q22,1)	Worthless_Assisting (Q22,7)	Early-Late (Q21,1)	Small-Large (Q21,2)
Intuitiveness_NYS (Q17)	1.000	0.996	0.964	0.985	0.990	0.968	0.980	0.979	0.879	0.938	0.926	0.975	0.963	0.986	0.980	0.979	-0.244	-0.460
Convincingness_NYS (Q18)		1.000	0.977	0.990	0.994	0.979	0.984	0.983	0.900	0.934	0.922	0.975	0.971	0.989	0.981	0.982	-0.270	-0.423
Unattractive-Attractive (Q21,4)			1.000	0.983	0.988	0.988	0.985	0.984	0.919	0.920	0.913	0.956	0.968	0.982	0.972	0.974	-0.251	-0.347
Bad-Good (Q22,3)				1.000	0.998	0.987	0.988	0.987	0.919	0.956	0.943	0.983	0.987	0.997	0.991	0.992	-0.231	-0.397
Undesirable-Desirable (Q22,8)					1.000	0.990	0.992	0.992	0.910	0.948	0.936	0.980	0.981	0.996	0.989	0.989	-0.218	-0.415
Annoying_Nice (Q22,4)						1.000	0.995	0.996	0.876	0.906	0.888	0.949	0.957	0.977	0.961	0.963	-0.175	-0.447
Unpleasant_Pleasant (Q22,2)							1.000	0.999	0.866	0.909	0.893	0.954	0.954	0.980	0.966	0.967	-0.206	-0.474
Irritating-Likeable (Q22,6)								1.000	0.862	0.911	0.894	0.956	0.952	0.980	0.966	0.965	-0.194	-0.489
SleepInducing_RaisingAlertness (Q22,9)									1.000	0.939	0.950	0.930	0.964	0.934	0.940	0.950	-0.407	-0.023
Intuitiveness_YS (Q19)										1.000	0.997	0.989	0.978	0.970	0.983	0.979	-0.238	-0.296
Convincingness_YS (Q20)											1.000	0.980	0.974	0.962	0.977	0.974	-0.265	-0.237
Unclear-Clear (Q21,3)												1.000	0.986	0.992	0.996	0.992	-0.258	-0.377
Superfluous_Effective (Q22, 5)													1.000	0.992	0.993	0.996	-0.266	-0.273
Composite Score														1.000	0.998	0.998	-0.252	-0.370
Useless_Useful (Q22,1)															1.000	0.998	-0.255	-0.352
Worthless_Assisting (Q22,7)																1.000	-0.276	-0.319
Early-Late (Q21,1)																	1.000	-0.318
Small-Large (Q21,2)																		1.000

Fig. 8. Correlation matrix for the means of the scores of the AR interfaces (n = 9). Responses to Q22 Items 1, 2, 4, 5, 7, 9 were reversed with respect to the questionnaire. The variables are sorted based on hierarchical clustering, i.e., similarity with the other variables.

Table 1
Means with standard deviations in parentheses for the composite scores (z-scores) ($n = 992$). Also shown are results for pairwise comparisons.

No	AR interface	Composite score	1	2	3	4	5	6	7	8	9
1	Augmented zebra crossing	0.32 (0.89)									
2	Planes on vehicle	−0.26 (1.01)	x								
3	Conspicuous looming planes	−0.35 (1.00)	x	x							
4	Field of safe travel	−0.12 (1.00)	x	x							
5	Fixed pedestrian lights	0.28 (0.88)		x	x	x					
6	Virtual fence	0.04 (1.00)	x	x	x	x	x				
7	Phantom car	−0.52 (1.05)	x	x	x	x	x	x			
8	Nudge HUD	0.37 (0.85)		x	x	x	x	x	x		
9	Pedestrian lights HUD	0.25 (0.86)		x	x	x		x	x	x	

Note. ‘x’ marks pairs of conditions that are statistically significantly different from each other, computed using paired-samples *t*-tests (*df* = 991).

(0.97), $t(981) = -1.71$, $p = 0.087$; *Nudge HUD*: Mean (*SD*) males/females: 0.37 (0.84)/0.37 (0.85), $t(981) = -0.06$, $p = 0.952$).

Country: The composite score of each interface was examined across the respondents' countries of residence (Fig. 9). The mean composite scores of the nine AR interfaces correlated again strongly. More specifically, for the 10 pairs of countries, correlations ranged between $r = 0.972$ (between Germany and Sweden) and $r = 0.992$ (between Norway and Sweden). A repeated-measures ANOVA of the composite score, with the AR interface as within-subject factor and country as between-subjects factor showed a significant effect of AR interface, $F(8, 7896) = 194.1$, $p < 0.001$, partial $\eta^2 = 0.16$, and no significant effect of country, $F(4, 987) = 0.82$, $p = 0.515$, partial $\eta^2 = 0.00$, and no significant AR interface × country interaction, $F(32, 7896) = 0.69$, $p = 0.902$, partial $\eta^2 = 0.00$.

Age: A repeated-measures ANOVA of the composite score, with the AR interface as a within-subject factor and age (45 or younger vs 46 or older) as a between-subjects factor showed a significant effect of AR interface, $F(8, 7920) = 195.2$, $p < 0.001$, partial $\eta^2 = 0.16$, and no significant effect of age group, $F(1, 990) = 0.44$, $p = 0.506$, partial $\eta^2 = 0.00$, and no significant AR interface × age group interaction, $F(8, 7920) = 1.52$, $p = 0.143$, partial $\eta^2 = 0.00$. The corresponding scatter plot is found in the supplementary material (Fig. S3).

Education: A repeated-measures ANOVA of the composite score, with the AR interface as a within-subject factor and educational attainment (university degree, trade/technical/vocational training, none of these) as a between-subjects factor showed a significant effect of AR interface, $F(8, 7912) = 167.8$, $p < 0.001$, partial $\eta^2 = 0.15$, and no significant effect of education, $F(2, 989) = 0.72$, $p = 0.489$, partial $\eta^2 = 0.00$, and no significant AR interface × education interaction, $F(16, 7912) = 0.98$, $p = 0.476$, partial $\eta^2 = 0.00$. The corresponding scatter plots are found in the supplementary material (Figs. S4 and S5).

It is noteworthy that although the overall composite score (i.e., averaged across the nine AR interfaces) did not correlate significantly with gender ($r = 0.01$ [1 = male, 2 = female]), age ($r = 0.02$), the highest level of education completed ($r = 0.04$, [1 = university degree, 2 = trade/technical/vocational training, 3 = none of these]), having ever used a VR headset (Q7; $r = -0.01$ [1 = no, 2 = yes]), or having ever used AR apps or games (Q8; $r = 0.02$ [1 = no, 2 = yes]), it did correlate moderately with willingness to use AR glasses ($r = 0.33$, 0.32, and 0.35 for Q9, Q10, and Q11, respectively) and with the ATI scale of technology affinity (Q6; $r = 0.22$). It is also noteworthy that older participants were less likely to have ever used VR (Q7; $r = -0.30$) or AR (Q8; $r = -0.44$, respectively).

3.4. Textual responses (Q23)

An average of 46 comments were extracted per interface. The subset of comments was further filtered down to retain four informative comments per concept, split equally between positive and negative valence (Table 2). These final selected comments were deemed representative of some of the major themes that arose per concept.

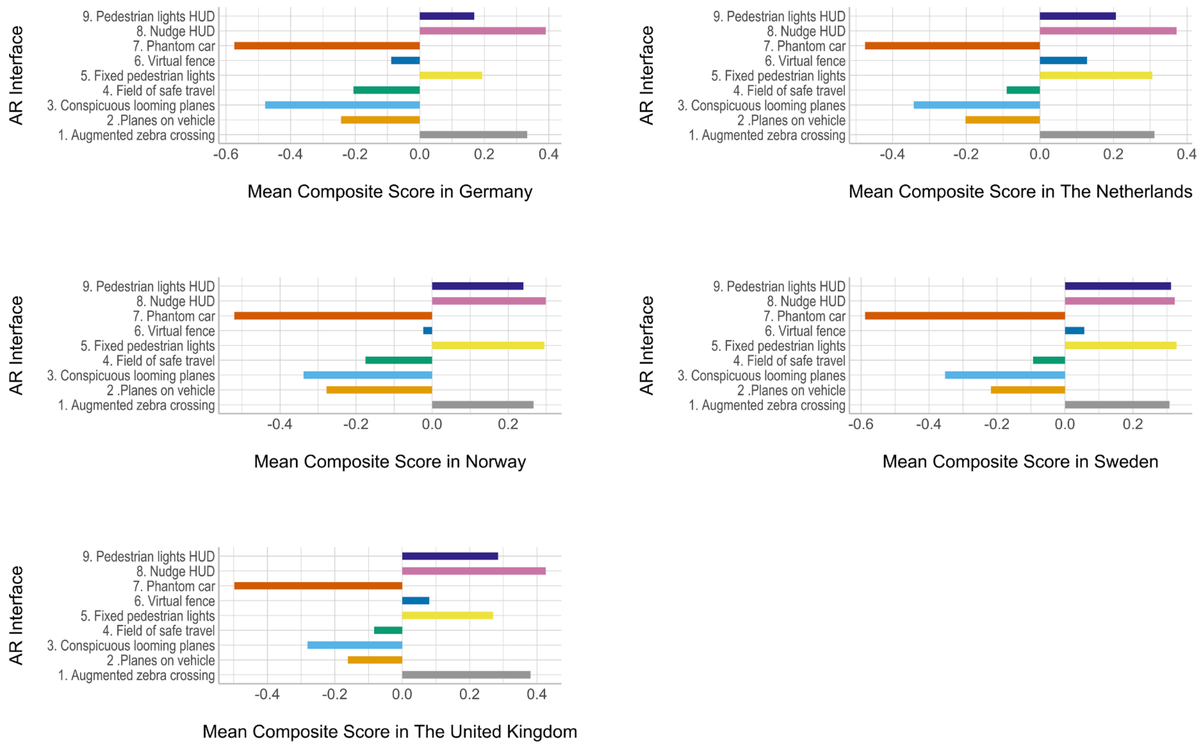


Fig. 9. Bar plots of the composite score of each interface per respondents' country. The standard deviation across respondents for the 45 depicted AR interface × country combinations ranges between 0.78 and 1.12.

Table 2
Sample of four comments per interface, split based on positive or negative sentiment. Spelling and grammar mistakes were not corrected.

AR Interface	Positive Comments	Negative Comments
Augmented zebra crossing	“A good idea. The zebra crossing is familiar to every-one. The big red cross over the crossing should make it clear not to cross.”	“It’s clear what the images mean but it doesn’t fill me with confidence regarding when it would be safe to cross the road. I think if you are not looking at the approaching vehicle you will always be in danger because you are not aware as to what it is doing, moving or stopping.”
	“Very clear and presumably understandable by most people including children once the different colours are explained to them.”	“The video signalling do not cross the road, I think is very clear. However, the video signalling that it is safe to cross is not so clear. The green lines either side of the pedestrian crossing did not immediately make me think it was safe, a green tick symbol maybe would’ve been better.”
Planes on vehicle	“[C]orrect colours for alert and safeness.”	“[T]he walking man on the green background made sense but the hand on the red background was unclear. i didn’t like it moving with the car, would prefer it to be in your face […]”
	“[B]etter variant because the size stays the same and symbols are clearer.”	“The problem with this signal, is that it just signals something about the car, not about the pedestrians”.
Conspicuous looming planes	“Very effective, the colour and hand signal stands out well.”	“[T]he colours are still very clear to understand: red for warning and green for no danger BUT as the vehicle approaches from the right side (around the corner) it was difficult so identify the signs written on the coloured boxes, it was kind of a weird perspective and therefore irritating. [A]s the stop/go signs where moving with the car and where not “fixed” at the top of my AR glasses, I had to think twice if these instructions were meant for me as a pedestrian or if there was another issues not concerning me.”
	“[T]he warning one was much better than the yielding one as the logo became larger as potential danger increased. [T]he change in size of the yielding one was hardly noticeable.”	“I wondered when something would actually appear in the screen. It took forever before I realised the notification was actually on the car itself. I find this visualisation absolutely useless.”
Field of safe travel	“I think it is somewhat useful as it shows the path of the vehicle.”	“[T]he green corridor has me confused, you see the car coming, with a corridor ahead, that makes me think it will drive on instead of stop.”
	“There was good warning time to let me know whether I was to cross or not. I also liked how the red and green showed up a good distance off too. Very clear.”	“The beam in the ‘stop video’ looks more like a red carpet, which I guess is something every-one would like to walk on.”
Fixed pedestrian lights	“This interface has been familiar and useful to me for as long as I remember, using it is highly intuitive and I see no need to alter it.”	“I think the sign for triggered too late for the non-yielding state, which would be more of a problem as I might already have started my journey across the street which can be a risk if the vehicle expects pedestrians to stand still. Otherwise the sign with a pole is very much familiar to me in my cultural context and therefore easily understood.”
	“[T]he interface is very clear/understandable as traffic lights are common in everyday lifeit includes people who are not able to readit seems like a ‘no energy’ interaction for me as I already know everything I need to know and do not have to think about it.”	“The signals are good, but optically too small and might well be overseen depending on the device holder (age, sight) or the background (lots of distraction on the street).”
Virtual fence	“It creates a safe feeling by creating a virtual wall.”	“I like the crossing part of this as previously stated, but pairing it with walls is really confusing. When you just see the red one, you immediately think they are walls to stop the car from going through and it looks like you are being given access through the crossing. The green one is better, but together confusing.”
	“Very clear in terms of the obvious colour difference but also in the size of the warnings. Very functional!”	“I realised in all examples so far, I enjoy the green signs more. I found this red one being wayyyyy too big and it literally made me jump when it appeared. It was also not clear to me that it signalled do not cross, except the red colour. When I could compare it with the green sign which was more intuitive it was clear that red meant stop. Before that I saw the red more as a frame/hallway around the zebra crossing.”
Phantom car	“[T]he phantom was very fast and clear and really did signal the options I had its sustainable as well.”	“[D]on’t like the look. reminds me of a video game. so I guess it can be dangerous cause you feel like in a game.”
	“Really good looking and easily understandable.”	“[T]he trouble is it’s just a bit too attractive and your brain does what it always does when you see something really attractive (particularly cars) and it goes ’WOW!’ When it does that it sort of sucks up all of your attention and you actually pay less attention to the other car. You almost forget about it.”
Nudge HUD	“I liked this one. People are pretty used to something similar to a notification like this and the colour + text makes it even easier to understand it.”	“[…] I feel the non-yielding state should specify ’do not cross’ as opposed to just stating a vehicle is approaching. The yielding state clearly states safe to cross so the message is much clearer with no room for misinterpretation.”

(*continued on next page*)

Table 2 (*continued*)

AR Interface	Positive Comments	Negative Comments
Pedestrian lights HUD	"This again empowers the user to make a choice based on their actions, not based on what the car is doing. It is much bigger then some, but in some ways less distracting. More functional."	"This example is clear enough, but a busy road is not like this. Except of cars, it can be running pets, pedestrians, bicycles coming from behind… It is dangerous to rely on this system, I think."
	"The best so far because you get the information in the same direction so you are looking for incoming traffic. Very nice."	"This is a lot clearer since it already relies on traffic rules that are now established in our society. I still have the feeling though that even if it is green that you would hold back a little bit with crossing the road since the car drives pretty fast towards you and I would only cross the street if the car is completely still."
	"Because the interface uses an image that I am already acquainted with (as are most members of the general public, including children and senior citizens) I found it to be very effective in indicating to me whether I could or could not cross the road safely."	"The image is clearly recognisable as one which indicates whether or not to cross. My only concern is that it is too small. It actually took me a few seconds to work out where it was. It could, of course,be that in time users would automatically focus on that part of their vision, and see the signal, but for this test, I found it worrying."

3.5. *Preferred AR interfaces and use of augmented reality in traffic*

The results of the final questionnaire section (Table 3) showed that 66 % of respondents felt that communication using AR interfaces in future traffic would be useful (Q24). Furthermore, 72 % preferred interfaces mapped to the street over those on the vehicle (Q25), 52 % preferred interfaces that included text rather than just graphical elements (Q26), and 51 % preferred head-locked over world-locked interfaces (Q27). Moreover, 62 % would like to have the ability to customise the AR interfaces (Q28), and 47 % indicated they would likely use AR interfaces as an aid for crossing in front of vehicles if they owned AR glasses (Q29).

4. Discussion

An online questionnaire study, aiming to evaluate nine AR interfaces for pedestrian-vehicle interaction, resulted in 992 valid respondents. Respondents were asked to rate the interfaces, presented in videos, on several qualities such as intuitiveness, convincingness, aesthetics, usefulness, and satisfaction.

4.1. *Interface preference by respondents*

When considering the intuitiveness and convincingness ratings (Figs. 6 and 7), and the composite score in Table 1, it can be asserted that AR interfaces that incorporated traditional traffic elements (*Augmented zebra crossing, Fixed pedestrian lights*, and *Pedestrian lights HUD*) and those that were head-locked performed better than the others. In addition, respondents indicated their preference for head-locked interfaces in the final responses of the questionnaire (Table 3).

The 'genius' design approach yielded a number of AR interfaces that were theoretically interesting but flawed from a user's point of view. The findings can retrospectively be explained by legacy design principles, which some AR interfaces adhered to and others did not (see Wickens et al., 2004, for thirteen established principles of display design). For example, although the *Phantom car* was designed to adhere to the principle of predictive aiding (since it showed the future position of the car), and the *Field of safe travel* adhered to the principle of ecological interface design (Kadar & Shaw, 2000; Tabone, Lee, et al., 2021; Waldenström, 2011), these two interfaces may have failed to comply with other design principles, such as redundancy gain (these interfaces displayed a coloured element, but no redundant icon or text), the proximity compatibility principle (it may be hard to perceptually separate the *Phantom car* from the real car), and the principle of top-down processing (participants are likely unfamiliar with these concepts). The most

Table 3

Descriptive statistics (i.e., means (*M*), standard deviations (*SD*), and relative frequencies) for the final questions.

Question	*M*	*SD*	Relative Frequencies				
			Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly Agree (5)
In future traffic, the communication from AR interfaces would be useful for crossing the road (Q24)	3.70	1.02	4.7 %	6.7 %	23.1 %	45.4 %	20.2 %
I prefer interfaces mapped to the street rather than on the vehicle (Q25)	3.98	0.98	1.9 %	6.1 %	19.5 %	37.3 %	35.2 %
I prefer interfaces with text rather than interfaces with just graphical elements (Q26)	3.44	1.09	5.3 %	14.0 %	28.5 %	35.4 %	16.7 %
I prefer interfaces that move around with my head rather than interfaces that stay fixed (Q27)	3.38	1.13	7.3 %	14.3 %	27.5 %	35.0 %	15.9 %
I would like to have the ability to customise these AR interfaces (Q28)	3.71	0.95	3.0 %	5.4 %	29.5 %	41.4 %	20.6 %
Now that I have seen these interfaces, if I own AR glasses, I am likely to use such interfaces as an aid for crossing in front of vehicles (Q29)	3.30	1.10	8.9 %	11.6 %	32.9 %	34.4 %	12.3 %

successful AR concepts, such as the *Augmented zebra crossing* and *Pedestrian lights* did adhere to the latter three principles, as described by Tabone, Lee, et al. (2021). The current observations also highlight the importance of involving the target user earlier on in the process through the use of a user-centred design methodology (Gulliksen et al., 2003) and to not rely on genius design only. The involvement of the target user early in the process could be achieved through focus groups, interviews, and card sorting, among other methods (Norman, 2013).

On the technical side, it is to be noted that the different AR interfaces involve different sensor and computational requirements (for an overview, see Tabone, Lee, et al., 2021). For instance, AR interfaces presented on the AV itself would have to rely on computer-vision techniques on the pedestrian's side, or vehicle-to-pedestrian communication of the AV's position and speed. The nudge interfaces, however, are considerably simpler and would only require the wireless communication of the AV's stopping intent to the pedestrian. These different sensor requirements were not presented to the respondents, nor were they considered in the evaluation of the AR interfaces.

Additionally, our study found that the means of questionnaire items were very strongly correlated and that the 15 acceptance-related items, in the aggregate, were well-represented by a single composite score. A recommendation that follows is that future research into the population-level mean acceptance of HMI concepts could just as well use a single acceptance item (such as a five-point scale ranging from bad to good) instead of multiple acceptance-related items. This finding aligns with previous research on the acceptance of automated driving systems, which indicated that different acceptance dimensions are hardly distinguishable and that a single factor of acceptance provides a better representation of the data (De Winter & Nordhoff, 2022; Nees & Zhang, 2020).

There were, however, two items that did not correlate strongly with the composite score, namely items related to the physical parameters of interface size and timing. As shown in Table S1 in the Supplementary Material, all nine AR interfaces yielded equivalent ratings (between 2.91 and 3.12) on the scale from 1 (too early) to 5 (too late) (Q21, Item 1), which can be explained by the fact that all interfaces were triggered at the same moment in the video. The small differences may have been caused by proximity (e.g., *Field of safe travel* extends in front of the car, "a sort of tongue protruding forward along the road"; Gibson & Crooks, 1938, p. 454), which might give participants the illusion that the interface was triggered early. The size ratings (Q21, Item 2) were also close to the midpoint for the nine interfaces, i.e., between 2.56 for the *Pedestrian lights* HUD and 3.37 for the *Virtual fence*. The differences in perceived size can also be explained by the actual size of the interfaces (see Fig. 2).

In the aggregate, different groups of participants reached similar conclusions on what they deemed to be a 'good' interface, i.e., results were similar regardless of gender, age, or country. Anecdotally, it is often believed that there are major cultural differences among pedestrians in that an eHMI that is found to work well in one country may not be received well in another country (see quotes of Bärgman, Hagenzieker, Krems and Ackerman, and Stanton in Tabone, De Winter, et al., 2021). The results of the present study suggest that these cultural differences are less strong as may be believed, at least for the five European countries under investigation. Our findings mirror those of others (Bazilinskyy et al., 2019; Singer et al., 2022) who found cross-cultural robustness of eHMIs in a larger number of countries on different continents.

While the online questionnaire was generally well distributed across the set quotas, it should be noted that the represented countries of residence were exclusively Western and Northern European. Therefore, cultural differences may have been relatively small. Several studies reported differences between the perceived clarity of eHMIs among participants from China versus Western Europe (Joisten et al., 2021; Lanzer et al., 2020; Weber et al., 2019). Whether or not cultural differences become apparent may depend on the clarity of the task instructions in the experiment and participants' prior expectations rather than the eHMI content itself, as noted by Singer et al. (2022).

4.2. Free-text comments

The textual inputs and opinions of the respondents were varied. Some respondents reported that the interfaces on the road surface could distract pedestrians from approaching vehicles, while others considered interfaces on the vehicle a hazard, since these blocked the visibility of the oncoming vehicle. In a number of instances, respondents indicated that they preferred the non-yielding state over the yielding state, with the former being regarded as more clear and intuitive. In fact, the intuitiveness and convincingness, as shown in Fig. 7, tended to favor the non-yielding state, except for a number of interfaces (*Planes on vehicle, Virtual fence*). Respondents described the yielding state for *Virtual fence* as clearer, but some labelled the non-yielding state as dangerous because the presence of a zebra crossing might tempt pedestrians to cross irrespective of the red gate. Similarly, the non-yielding state for the *Field of safe travel* was labelled as potentially dangerous by some because it looked like a red carpet that invited them to walk on it.

Another prevalent theme was that some respondents felt that at times it was not clear to whom the communication referred, i.e., the pedestrian or the vehicle itself. For example, the hand symbol on the *Planes on vehicle* was sometimes misinterpreted as indicating a problem with the vehicle. The *Planes on vehicle* and *Conspicuous looming planes* interfaces, which project planes on the vehicle, drew concerns about a blocked view of the vehicle, yet at the same time, the looming planes concept was commended for its clarity in communicating danger. These observations reveal the issue of unintended effects resulting from 'genius designs', where the intention is not fully grasped by the user. Our findings resonate with broader issues in human factors, namely that "*the actual, rather than presumed, impact of new technology is usually quite surprising, unintended, and even counterproductive*" (Woods & Dekker, 2000, p. 276).

Similar to the observations derived from the statistical analysis, the interfaces based on more traditional traffic elements were labelled as more understandable and intuitive due to familiar symbology (e.g., zebra crossing, traffic light). The 'worst' performing interface (*Phantom car*), while commended for its aesthetic qualities, received various critical descriptions, such as 'confusing', 'frightening', 'scary', 'startling', 'spooky', and 'unclear'. In fact, some described the interface as a video game, which in a sense confirms the original design direction of the *Phantom car* concept from Tabone, Lee, et al. (2021), where the idea of ghost cars from racing video games was drawn upon.

The HUD interfaces were praised for being ‘logical’, ‘visible’, ‘clear’, and ‘perfect’ to capture the attention of distracted pedestrians. However, it was also stated that HUDs could be a distraction from other hazards, especially when text is used (for further discussion on text-based eHMIs, see Bazilinskyy et al., 2019). Moreover, a number of respondents complained that the text was in English, and that this would be a danger for pedestrians unfamiliar with the language. The latter feedback resonates with an advantage of AR communication, where personalization of the interface could solve the language issue. In fact, 62 % of the respondents were in favour of such a possibility. Finally, a number of times, respondents suggested that they would still rely on the vehicle coming to a full stop before making any decision, confirming that implicit communication plays an important role in shaping pedestrian decisions (Lee et al., 2021).

4.3. Limitations and future work

Although the online questionnaire was distributed to a wide respondent pool, the analysis revealed that more than half of the respondents (54 %) reported having attained a university degree. Research suggests that university graduates are more inclined towards the adoption and usage of technology (Burton-Jones & Hubona, 2005; Nielsen & Haustein, 2018). At the same time, we found strong convergence in ratings for participants with and without a university degree, suggesting that educational level was not an important moderator of the current findings (see Figs. S4 and S5). A possible reason is that participants were not asked to understand or use complex technology; instead, the present task was largely-one of perceptual nature.

A further limitation is that the high correlation of acceptance-related items may have arisen from the uniform questionnaire format, giving rise to acquiescence bias. However, this limitation may not be significant as the acceptance scale (Q22) contained reversed items (from high to low, and from low to high), yet these items still correlated very strongly with the responses to the intuitiveness and convincingness items.

A number of free-text comments mentioned drivers being blinded by the interfaces that appeared on the car, indicating that those respondents did not fully grasp what AR technology is. Additionally, there were instances where the terms ‘AR’ and ‘VR’ were used interchangeably in the comments, with a number of respondents expressing total opposition towards wearing ‘VR headsets’ when they walk around outside. This confusion may have been caused by the fact that participants only saw VR videos of AR concepts, rather than experiencing AR themselves. That said, such confusion would only have affected the overall understanding of AR, and probably not the relative differences in the participants’ assessments of the nine AR concepts.

Many respondent comments were unusable in the thematic analysis. While gibberish text entries were uncommon, many of the textual comments were too brief to provide useful information (e.g., “This one was clear”). This highlights a limitation of online studies, where there is the risk that some respondents do not thoroughly read the information provided at the beginning or aim to complete the questionnaire items quickly. A further limitation of online studies with videos is that, while offering high repeatability, they do not offer high ecological validity and present only low perceived risk to participants (for a similar discussion, see Fuest et al., 2020; Petzoldt et al., 2018; Tabone, De Winter, et al., 2021).

A further limitation was that the environment consisted of a one-way road, with only one vehicle. The addition of more traffic, with varying trajectories, would add more natural cues to the testing environment. It can be hypothesized that the *Nudge HUD* will be particularly effective when multiple vehicles approach from different directions since the *Nudge HUD* does not require the pedestrian to distribute attention across those vehicles. In contrast, the *Planes on vehicle* require the pedestrian to first locate the planes in the environment before crossing, which may be time-consuming and inefficient. A potential advantage of *Planes on vehicle*, on the other hand, is that it may prevent overreliance in situations of e.g., vehicle-to-pedestrian communication failure. Another limitation was the lack of environmental sound, and the fact that participants were not asked to interact with the scene (e.g., to indicate when it is safe to cross). To better understand the behaviour of users of such interfaces, ecological validity must be increased. Therefore, in the future, the stimuli could be presented to the participants in a virtual simulation environment and ultimately, in the real world.

6. Conclusion

Nine augmented reality interfaces for pedestrian-vehicle interaction were presented in a video-based online study that yielded 992 respondents from Germany, the Netherlands, Norway, Sweden, and the United Kingdom. Each interface was shown in its non-yielding and yielding states at a pedestrian crossing area represented in a VR environment. Respondents were asked to rate each interface based on its intuitiveness and convincingness in communicating whether or not a vehicle would yield. Other ratings related to functional and aesthetic qualities, usefulness, and satisfaction.

Statistical and qualitative thematic analysis indicated that respondents preferred head-locked interfaces over their world-locked counterparts, with interfaces employing traditional traffic interface elements receiving higher ratings than others. These results indicated that legacy design principles performed better than designs generated through an expert-based approach (‘genius’ design), further highlighting the importance of involving the user early in the process. A further qualitative analysis provided more context to the ratings, such as the preference of the non-yielding state over the yielding state for a number of interfaces, preference towards traditional traffic symbols, and reliance on implicit cues.

Responses related to the general use of interfaces indicated a preference for interfaces that are mapped to the street instead of the vehicle. Moreover, respondents preferred interfaces that make use of text compared to interfaces that use just graphical elements, and interfaces that are head-locked rather than world-locked. Most respondents also indicated that they would like to personalise the AR interfaces, and that communication using AR interfaces in future traffic would be useful.

Although the current online study offered an indication of what kinds of AR interfaces, placement in the world, and design elements are more suitable for pedestrian-vehicle interactions, there are limitations related to the ecological validity dimension of the study. In order to better understand the behaviour of potential users of the system, in the future, the ecological validity of such a user evaluation should be increased.

The practical implications of the present study depend on the progression in vehicle automation and communication, and in AR. It seems plausible that computers will become increasingly compact, and that the use of AR, either via handheld or head-mounted devices will become increasingly feasible in the real world. At the same time, questions about inclusivity, affordability, and user acceptance remain to be addressed, as discussed by Tabone, De Winter, et al. (2021). A likely way forward is that the use of AR for pedestrians will see its introduction first in professional transportation contexts (e.g., warehouses, airport personnel) before becoming available to the general public.

CRediT authorship contribution statement

Wilbert Tabone: Conceptualization, Data curation, Investigation, Formal analysis, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing. **Riender Happee:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Jorge García:** Resources, Software. **Yee Mun Lee:** Conceptualization, Methodology. **Maria Luce Lupetti:** Conceptualization. **Natasha Merat:** Conceptualization, Methodology. **Joost de Winter:** Conceptualization, Investigation, Formal analysis, Methodology, Validation, Writing – original draft, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

A link to the data has been included in the manuscript file.

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.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trf.2023.02.005.

References

Ackermans, S., Dey, D., Ruijten, P., Cuijpers, R. H., & Pfleging, B. (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*. https://doi.org/10.1145/3313831.3376197

Adell, E. (2010). Acceptance of driver support systems. In *Proceedings of the European conference on human centred design for intelligent transport systems*, Berlin, Germany (pp. 475–486).

Bai, S., Legge, D. D., Young, A., Bao, S., & Zhou, F. (2021). Investigating external interaction modality and design between automated vehicles and pedestrians at crossings. In *Proceedings of the 2021 IEEE international intelligent transportation systems conference (ITSC)*. https://doi.org/10.1109/ITSC48978.2021.9564867

Bazilinskyy, P., Dodou, D., & De Winter, J. (2019). Survey on eHMI concepts: The effect of text, color, and perspective. *Transportation Research Part F: Traffic Psychology and Behaviour, 67*, 175–194. https://doi.org/10.1016/j.trf.2019.10.013

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*, Article 103450. https://doi.org/10.1016/j.apergo.2021.103450

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., … Johnson, V. E. (2018). Redefine statistical significance. *Nature Human. Behaviour, 2*, 6–10. https://doi.org/10.1038/s41562-017-0189-z

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'? In *IEEE international conference on systems, man and cybernetics (SMC)* (pp. 3721–3728), Toronto, Canada. https://doi.org/10.1109/SMC42975.2020.9282998.

Bhorkar, G. (2017). *A survey of augmented reality navigation.* arXiv. https://arxiv.org/abs/1708.05006.

Bindschädel, J., Krems, I., & Kiesel, A. (2022). Active vehicle pitch motion for communication in automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 87*, 279–294. https://doi.org/10.1016/j.trf.2022.04.011

Buker, T. J., Vincenzi, D. A., & Deaton, J. E. (2012). The effect of apparent latency on simulator sickness while using a see-through helmet-mounted display: Reducing apparent latency with predictive compensation. *Human Factors, 54*, 235–249. https://doi.org/10.1177/0018720811428734

Burton-Jones, A., & Hubona, G. S. (2005). Individual differences and usage behavior: Revisiting a technology acceptance model assumption. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 36*, 58–77. https://doi.org/10.1145/1066149.1066155

Calvi, A., D'Amico, F., Ferrante, C., & Ciampoli, L. B. (2020). Effectiveness of augmented reality warnings on driving behaviour whilst approaching pedestrian crossings: A driving simulator study. *Accident Analysis & Prevention, 147*, Article 105760. https://doi.org/10.1016/j.aap.2020.105760

Carmona, J., Guindel, C., Garcia, F., & De la Escalera, A. (2021). eHMI: Review and guidelines for deployment on autonomous vehicles. *Sensors, 21*, 2912. https://doi.org/10.3390/s21092912

Cavallo, V., Dommes, A., Dang, N. T., & Vienne, F. (2019). A street-crossing simulator for studying and training pedestrians. *Transportation Research Part F: Traffic Psychology and Behaviour, 61*, 217–228. https://doi.org/10.1016/j.trf.2017.04.012

Colley, M., Eder, B., Rixen, J. O., & Rukzio, E. (2021). Effects of semantic segmentation visualization on trust, situation awareness, and cognitive load in highly automated vehicles. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. https://doi.org/10.1145/3411764.3445351

Currano, R., Park, S. Y., Moore, D. J., Lyons, K., & Sirkin, D. (2021). Little road driving hud: Heads-up display complexity influences drivers' perceptions of automated vehicles. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, Yokohama Japan. https://doi.org/10.1145/3411764.3445575.

Dancu, A., Vechev, V., Ünlüer, A. A., Nilson, S., Nygren, O., Eliasson, S., Barjonet, J.-E., Marshall, J., & Fjeld, M. (2015). Gesture bike: Examining projection surfaces and turn signal systems for urban cycling. In *Proceedings of the 2015 international conference on interactive tabletops & surfaces* (pp. 151–159). https://doi.org/10.1145/2817721.2817748

De Winter, J., & Dodou, D. (2022). External human-machine interfaces: Gimmick or necessity. *Transportation Research Interdisciplinary Perspectives, 15*, Article 100643. https://doi.org/10.1016/j.trip.2022.100643

De Winter, J. C. F., & Hancock, P. A. (2015). Reflections on the 1951 Fitts list: Do humans believe now that machines surpass them? In N. V. Las Vegas (Ed.), *Proceedings of the 6th international conference on applied human factors and ergonomics (AHFE)* (pp. 5334–5341). https://doi.org/10.1016/j.promfg.2015.07.641

De Winter, J., & Nordhoff, S. (2022). Acceptance of conditionally automated cars: Just one factor? *Transportation Research Interdisciplinary Perspectives, 15*, Article 100645. https://doi.org/10.1016/j.trip.2022.100645

Deb, S., Strawderman, L., Carruth, D. W., DuBien, J., Smith, B., & Garrison, T. M. (2017). Development and validation of a questionnaire to assess pedestrian receptivity toward fully autonomous vehicles. *Transportation Research Part C: Emerging Technologies, 84*, 178–195. https://doi.org/10.1016/j.trc.2017.08.029

Dey, D., & Terken, J. (2017). Pedestrian interaction with vehicles: roles of explicit and implicit communication. In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications*, Oldenburg, Germany, 109–113. https://doi.org/10.1145/3122986.3123009.

Dey, D., Habibovic, A., Pfleging, B., Martens, M., & Terken, J. (2020b). 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*, Honolulu, HI. https://doi.org/10.1145/3313831.3376325.

Dey, D., Habibovic, A., Löcken, A., Wintersberger, P., Pfleging, B., Riener, A., Martens, M., & Terken, J. (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*, Article 100174. https://doi.org/10.1016/j.trip.2020.100174

Dietrich, A., Tondera, M., & Bengler, K. (2020). Automated vehicles in urban traffic: The effect of kinematics and eHMIs on pedestrian crossing behavior. *Advances in Transportation Studies, 2020*, 73–84.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation, 14*, 20. https://doi.org/10.7275/da8t-4g52

Dong, W., Wu, Y., Qin, T., Bian, X., Zhao, Y., He, Y., Xu, Y., & Yu, C. (2021). What is the difference between augmented reality and 2D navigation electronic maps in pedestrian wayfinding? *Cartography and Geographic Information Science, 48*, 225–240. https://doi.org/10.1080/15230406.2021.1871646

Endsley, T. C., Sprehn, K. A., Brill, R. M., Ryan, K. J., Vincent, E. C., & Martin, J. M. (2017). Augmented reality design heuristics: Designing for dynamic interactions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*, 2100–2104. https://doi.org/10.1177/154193121360200

Faas, S. M., Mathis, L.-A., & Baumann, M. (2020). External HMI for self-driving vehicles: Which information shall be displayed? *Transportation Research Part F: Traffic Psychology and Behaviour, 68*, 171–186. https://doi.org/10.1016/j.trf.2019.12.009

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human-Computer Interaction, 35*, 456–467. https://doi.org/10.1080/10447318.2018.1456150

Fuest, T., Michalowski, L., Träris, L., Bellem, H., & Bengler, K. (2018). Using the driving behavior of an automated vehicle to communicate intentions – a Wizard of Oz study. In *Proceedings of the 2018 21st international conference on intelligent transportation systems (ITSC)*, Maui, HI (pp. 3596–3601). https://doi.org/10.1109/ITSC.2018.8569486.

Fuest, T., Schmidt, E., & Bengler, K. (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*, 291. https://doi.org/10.3390/info11060291

Gibson, J. J., & Crooks, L. E. (1938). A theoretical field-analysis of automobile-driving. *The American Journal of Psychology, 51*, 453–471. https://doi.org/10.2307/1416145

Ginters, E. (2019). Augmented reality use for cycling quality improvement. *Procedia Computer Science, 149*, 167–176. https://doi.org/10.1016/j.procs.2019.01.120

Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., & Cajander, Å. (2003). Key principles for user-centred systems design. *Behaviour & Information Technology, 22*, 397–409. https://doi.org/10.1080/01449290310001624329

Hagenzieker, M., Boersma, R., Velasco, P. N., Ozturker, M., Zubin, I., & Heikoop, D. (2020). *Automated buses in Europe: An inventory of pilots*. Delft University of Technology. Technical Report.

Hensch, A.-C., Neumann, I., Beggiato, M., Halama, J., & Krems, J. F. (2019). Effects of a light-based communication approach as an external HMI for Automated Vehicles — A Wizard-of-Oz study. *Transactions on Transport Sciences, 10*, 18–32. https://doi.org/10.5507/tots.2019.012

Lau, M., Le, D. H., & Oehl, M. (2021). Design of external human-machine interfaces for different automated vehicle types for the interaction with pedestrians on a shared space. In: N. L. Black, W. P. Neumann, & I. Noy (Eds.), *Proceedings of the 21st congress of the international ergonomics association (IEA 2021)* (pp. 710–717). Cham: Springer. https://doi.org/10.1007/978-3-030-74608-7_87.

Hesenius, M., Börsting, I., Meyer, O., & Gruhn, V. (2018). Don't panic! Guiding pedestrians in autonomous traffic with augmented reality. In *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services adjunct*, Barcelona, Spain (pp. 261–268). https://doi.org/10.1145/3236112.3236148.

Innofact. (2022). INNOFACT AG Das MarktVORSPRUNGSinstitut. https://www.innofact.com.

Ishihara, S. (1917). *Tests for colour-blindness*. Tokyo: Kanehara Shuppan Co., Ltd.

Joisten, P., Liu, Z., Theobald, N., Webler, A., & Abendroth, B. (2021). Communication of automated vehicles and pedestrian groups: An intercultural study on pedestrians' street crossing decisions. In *Proceedings of MuC '21: mensch und computer 2021*, Ingolstadt, Germany (pp. 49–53). https://doi.org/10.1145/3473856.3474004.

Kadar, E. E., & Shaw, R. E. (2000). Toward an ecological field theory of perceptual control of locomotion. *Ecological Psychology, 12*, 141–180. https://doi.org/10.1207/S15326969ECO1202_02

Kaleefathullah, A. A., Merat, N., Lee, Y. M., Eisma, Y. B., Madigan, R., Garcia, J., & De Winter, J. C. F. (2020). External human–machine interfaces can be misleading: An examination of trust development and misuse in a CAVE-based pedestrian simulation environment. *Human Factors*. https://doi.org/10.1177/0018720820970751

Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher, 42*, 846–854. https://doi.org/10.1080/0142159X.2020.1755030

Kim, H., Gabbard, J. L., Anon, A. M., & Misu, T. (2018). Driver behavior and performance with augmented reality pedestrian collision warning: An outdoor user study. *IEEE Transactions on Visualization and Computer Graphics, 24*, 1515–1524. https://doi.org/10.1109/TVCG.2018.2793680

Lanzer, M., Babel, F., Yan, F., Zhang, B., You, F., Wang, J., & Baumann, M. (2020). Designing communication strategies of autonomous vehicles with pedestrians: an intercultural study. In *Proceedings of the 12th international conference on automotive user interfaces and interactive vehicular applications*, Virtual Event (pp. 122–131). https://doi.org/10.1145/3409120.3410653.

Lee, Y. M., Madigan, R., Giles, O., Garach-Morcillo, L., Markkula, G., Fox, C., Camara, F., Rothmueller, M., Vendelbo-Larsen, S. A., Rasmussen, P. H., Dietrich, A., Nathanael, D., Portouli, V., Schieben, A., & Merat, N. (2021). Road users rarely use explicit communication when interacting in today's traffic: Implications for automated vehicles. *Cognition, Technology & Work, 23*, 367–380. https://doi.org/10.1007/s10111-020-00635-y

Li, Y., Kamalasanan, V., Batista, M., & Sester, M. (2022). *Improving pedestrian priority via grouping and virtual lanes.* arXiv. https://arxiv.org/abs/2205.08783.

Löcken, A., Golling, C., & Riener, A. (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, The Netherlands (pp. 262–274). https://doi.org/10.1145/3342197.3344544.

Maruhn, P., Dietrich, A., Prasch, L., & Schneider, S. (2020). Analyzing pedestrian behavior in augmented reality—proof of concept. In *Proceedings of the 2020 IEEE conference on virtual reality and 3D user interfaces (VR)*, Atlanta, GA (pp. 313–321). https://doi.org/10.1109/VR46266.2020.00051.

Matviienko, A., Müller, F., Schön, D., Seesemann, P., Günther, S., & Mühlhäuser, M. (2022). BikeAR: Understanding cyclists' crossing decision-making at uncontrolled intersections using Augmented Reality. In *Proceedings of the CHI conference on human factors in computing systems*, New Orleans, LA. https://doi.org/10.1145/3491102.3517560.

Métayer, N., & Coeugnet, S. (2021). Improving the experience in the pedestrian's interaction with an autonomous vehicle: An ergonomic comparison of external HMI. *Applied Ergonomics, 96*, Article 103478. https://doi.org/10.1016/j.apergo.2021.103478

Mok, C. S., Bazilinskyy, P., & De Winter, J. C. F. (2022). Stopping by looking: A driver-pedestrian interaction study in a coupled simulator using head-mounted displays with eye-tracking. *Applied Ergonomics, 105*, Article 103825. https://doi.org/10.1016/j.apergo.2022.103825

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4*, 61–64. 10.20982/tqmp.04.2.p061.

Nees, M., & Zhang, J. (2020). *Acceptance of highly automated vehicles: A factor analysis approach.* PsyArXiv. https://doi.org/10.31234/osf.io/9qcjp.

Nielsen, T. A. S., & Haustein, S. (2018). On sceptics and enthusiasts: What are the expectations towards self-driving cars? *Transport Policy, 66*, 49–55. https://doi.org/10.1016/j.tranpol.2018.03.004

Nielsen, J. (2007). *The myth of the genius designer*. https://www.nngroup.com/articles/the-myth-of-the-genius-designer.

Nordhoff, S., Beuster, A., Kessel, T., Bjorvatn, A., Innamaa, S., Lehtonen, E., Malin, F., Madigan, R., Lee, Y.-M., Merat, N., & Louw, T. (2021). *Annual quantitative survey about user acceptance towards ADAS and vehicle automation* (Deliverable D7.1 of L3Pilot project). European Commission. https://l3pilot.eu/downloads#c81.

Norman, D. (2013). *The design of everyday things* (Revised and expanded edition.). Basic Books.

Onkhar, V., Bazilinskyy, P., Stapel, J. C. J., Dodou, D., Gavrila, D., & De Winter, J. C. F. (2021). Towards the detection of driver–pedestrian eye contact. *Pervasive and Mobile Computing, 76*, Article 101455. https://doi.org/10.1016/j.pmcj.2021.101455

Oudshoorn, M., De Winter, J. C. F., Bazilinskyy, P., & Dodou, D. (2021). Bio-inspired intent communication for automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour, 80*, 127–140. https://doi.org/10.1016/j.trf.2021.03.021

Petzoldt, T., Schleinitz, K., & Banse, R. (2018). Potential safety effects of a frontal brake light for motor vehicles. *IET Intelligent Transport Systems, 12*, 449–453. https://doi.org/10.1049/iet-its.2017.0321

Pichen, J., Yan, F., & Baumann, M. (2020). Towards a cooperative driver-vehicle interface: Enhancing drivers' perception of cyclists through Augmented Reality. In *Proceedings of the 2020 IEEE intelligent vehicles symposium (IV)*, Las Vegas, NV (pp. 1827–1832). https://doi.org/10.1109/IV47402.2020.9304621.

Pompigna, A., & Mauro, R. (2022). Smart roads: A state of the art of highways innovations in the Smart Age. *Engineering Science and Technology, an International Journal, 25*, Article 100986. https://doi.org/10.1016/j.jestch.2021.04.005

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. (2022). Qualtrics XM//The leading experience management software. https://www.qualtrics.com.

Rahman, M. M., Lesch, M. F., Horrey, W. J., & Strawderman, L. (2017). Assessing the utility of TAM, TPB, and UTAUT for advanced driver assistance systems. *Accident Analysis & Prevention, 108*, 361–373. https://doi.org/10.1016/j.aap.2017.09.011

Ranieri, M. (2020). *A new sense of direction with Live View*. https://blog.google/products/maps/new-sense-direction-live-view.

Rasouli, A., & Tsotsos, J. K. (2020). Autonomous vehicles that interact with pedestrians: A survey of theory and practice. *IEEE Transactions on Intelligent Transportation Systems, 21*, 900–918. https://doi.org/10.1109/TITS.2019.2901817

Rolland, J. P., Gibson, W., & Ariely, D. (1995). Towards quantifying depth and size perception in virtual environments. *Presence: Teleoperators and Virtual Environments, 4*, 24–49. https://doi.org/10.1162/pres.1995.4.1.24

Rouchitsas, A., & Alm, H. (2019). External human–machine interfaces for autonomous vehicle-to-pedestrian communication: A review of empirical work. *Frontiers in Psychology, 10*, 2757. https://doi.org/10.3389/fpsyg.2019.02757

Saffer, D. (2010). *Designing for interaction: Creating innovative applications and devices*. Berkeley, CA: New Riders.

Schieben, A., Wilbrink, M., Kettwich, C., Dodiya, J., Weber, F., Sorokin, L., Lee, Y.-M., Madigan, R., Markkula, G., Merat, N., Dietrich, A., & Kaup, M. (2019). Testing external HMI designs for automated vehicles–An overview on user study results from the EU project interACT. 9. Tagung Automatisiertes Fahren.

Singer, T., Kobbert, J., Zandi, B., & Khanh, T. Q. (2022). Displaying the driving state of automated vehicles to other road users: An international, virtual reality-based study as a first step for the harmonized regulations of novel signaling devices. *IEEE Transactions on Intelligent Transportation Systems, 23*, 2904–2918. https://doi.org/10.1109/TITS.2020.3032777

Sripada, A., Bazilinskyy, P., & De Winter, J. (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., 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. (2021). Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. *Transportation Research Interdisciplinary Perspectives, 9*, Article 100293. https://doi.org/10.1016/j.trip.2020.100293

Tabone, W., Lee, Y. M., Merat, N., Happee, R., & De Winter, J. C. F. (2021). Towards future pedestrian-vehicle interactions: Introducing theoretically-supported AR prototypes. In *Proceedings of the 13th international conference on automotive user interfaces and interactive vehicular applications*. https://doi.org/10.1145/3409118.3475149

Toh, C. K., Sanguesa, J. A., Cano, J. C., & Martinez, F. J. (2020). Advances in smart roads for future smart cities. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 476*, 20190439. https://doi.org/10.1098/rspa.2019.0439

Tong, Y., & Jia, B. (2019). An Augmented-Reality-based warning interface for pedestrians: User interface design and evaluation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 63*, 1834–1838. https://doi.org/10.1177/1071181319631413

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*, Article 866516. https://doi.org/10.3389/fcomp.2022.866516

Unity. (2022). Unity real-time development platform | 3D, 2D VR & AR engine. https://www.unity.com.

University of Leeds. (2022). Highly Immersive Kinematic Experimental Research (HIKER) pedestrian lab. https://uolds.leeds.ac.uk/facility/hiker-lab.

Uttley, J., Lee, Y. M., Madigan, R., & Merat, N. (2020). Road user interactions in a shared space setting: Priority and communication in a UK car park. *Transportation Research Part F: Traffic Psychology and Behaviour, 72*, 32–46. https://doi.org/10.1016/j.trf.2020.05.004

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

Volkswagen. (2020). From the luxury class to the compact segment: the augmented reality head-up display. https://www.volkswagen-newsroom.com/en/press-releases/from-the-luxury-class-to-the-compact-segment-the-augmented-reality-head-up-display-6730.

Von Sawitzky, T., Wintersberger, P., Löcken, A., Frison, A. K., & Riener, A. (2020). Augmentation concepts with HUDs for cyclists to improve road safety in shared spaces. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. https://doi.org/10.1145/3334480.3383022

Waldenström, C. (2011). Visualizing the field of safe travel increases performance in a naval movement task. In *Proceedings of the 2011 IEEE international multi-disciplinary conference on cognitive methods in situation awareness and decision support (CogSIMA), Miami Beach* (pp. 252–256). https://doi.org/10.1109/COGSIMA.2011.5753454

Wann, J. P., Rushton, S., & Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. *Vision Research, 35*, 2731–2736. https://doi.org/10.1016/0042-6989(95)00018-U

Weber, F., Chadowitz, R., Schmidt, K., Messerschmidt, J., & Fuest, T. (2019). Crossing the street across the globe: A study on the effects of eHMI on pedestrians in the US, Germany and China. In H. Krömker (Ed.), *HCI in mobility, transport, and automotive systems. HCII 2019* (pp. 515–530). Cham: Springer. https://doi.org/10.1007/978-3-030-22666-4_37.

Wickens, C. D., Gordon, S. E., & Liu, Y. (2004). *An introduction to human factors engineering.* Upper Saddle River, NJ: Pearson Prentice Hall.

Woods, D., & Dekker, S. (2000). Anticipating the effects of technological change: A new era of dynamics for human factors. *Theoretical Issues in Ergonomics Science, 1*, 272–282. https://doi.org/10.1080/14639220110037452