Evaluating Autonomous Vehicle External Communication Using a Multi-Pedestrian VR Simulator

TRAM THI MINH TRAN, Design Lab, The University of Sydney, Australia
CALLUM PARKER, Design Lab, The University of Sydney, Australia
XINYAN YU, Design Lab, The University of Sydney, Australia
DEBARGHA DEY, Information Science, Cornell Tech, USA
MARIEKE MARTENS, Department of Industrial Design, Eindhoven University of Technology, Netherlands
PAVLO BAZILINSKYY, Department of Industrial Design, Eindhoven University of Technology, Netherlands
MARTIN TOMITSCH, Transdisciplinary School, University of Technology Sydney, Australia

Fig. 1. The study setup with two pedestrian participants: (left) real-world view and (right) VR view.

With the rise of autonomous vehicles (AVs) in transportation, a pressing concern is their seamless integration into daily life. In multi-pedestrian settings, two challenges emerge: ensuring unambiguous communication to individual pedestrians via external Human–Machine Interfaces (eHMIs), and the influence of one pedestrian over another. We conducted an experiment (N=25) using a multi-pedestrian virtual reality simulator. Participants were paired and exposed to three distinct eHMI concepts: on the *vehicle*, within the surrounding *infrastructure*, and on the *pedestrian* themselves, against a baseline without any eHMI. Results indicate that all eHMI concepts improved clarity of communication over the baseline, but differences in their effectiveness were observed. While pedestrian and infrastructure communications often provided more direct clarity, vehicle-based cues at times introduced uncertainty elements. Furthermore, the study identified the role of co-located pedestrians: in the absence of clear AV communication, individuals frequently sought cues from their peers.

Authors' addresses: Tram Thi Minh Tran, tram.tran@sydney.edu.au, Design Lab, The University of Sydney, Sydney, Australia; Callum Parker, callum.parker@sydney.edu.au, Design Lab, The University of Sydney, Sydney, Australia; Xinyan Yu, xinyan.yu@sydney.edu.au, Design Lab, The University of Sydney, Sydney, Australia; Debargha Dey, debargha.dey@cornell.edu, Information Science, Cornell Tech, New York, USA; Marieke Martens, m.h.martens@tue.nl, Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands; Pavlo Bazilinskyy, p.bazilinskyy@tue.nl, Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands; Martin Tomitsch, Martin.Tomitsch@uts.edu.au, Transdisciplinary School, University of Technology Sydney, Sydney, Australia.

ACM 2474-9567/2024/9-ART130
https://doi.org/10.1145/3678506

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

Additional Key Words and Phrases: autonomous vehicles, external communication, eHMIs, vulnerable road users, vehicle-pedestrian interaction, scalability

ACM Reference Format:
Tram Thi Minh Tran, Callum Parker, Xinyan Yu, Debargha Dey, Marieke Martens, Pavlo Bazilinskyy, and Martin Tomitsch. 2024. Evaluating Autonomous Vehicle External Communication Using a Multi-Pedestrian VR Simulator. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 3, Article 130 (September 2024), 26 pages. https://doi.org/10.1145/3678506

1 INTRODUCTION

Despite the promise of increased road safety and efficiency with autonomous vehicles (AVs), their successful integration into everyday life depends on clear communication with pedestrians, among other factors. To bridge this gap, external Human-Machine Interfaces (eHMIs) have emerged to convey the intentions and operational states of AVs, extensively explored by both the research community and industry [16, 17, 54, 56]. Policymakers are also carefully drafting recommendations for effective rollout of eHMIs, with ongoing initiatives such as the SAE's J3134 standard [58] and the UNECE's recommendations [72]. In the most recent development, Mercedes has received approval to test turquoise-coloured marker lights for automated driving on California freeways [48]

While a considerable amount of eHMI research has focused on AV interactions with individual pedestrians, there is a notable scarcity of studies in multi-pedestrian contexts [14, 65]. The existing literature highlights a significant issue known as 'clarity of recipients' [17, 66], which contributes to pedestrians feeling uncertain [71], less safe [12], or mistakenly assuming that AV communications are directed specifically at them when they are not [19, 20]. This problem is particularly pronounced among co-located pedestrians—individuals who navigate alongside AVs without being part of a cohesive group [19]. Unlike pedestrians in groups, who benefit from coordinated behaviours that facilitate a unified response to AV communications, co-located pedestrians interpret these signals independently, and any resulting misunderstandings can potentially lead to unsafe situations. On-vehicle eHMIs, the most common form integrated into the AV's exterior, have been found to be limited in addressing these challenges within multi-pedestrian contexts [19].

In response to these limitations, researchers have proposed enhancing AV communication clarity by incorporating infrastructure and pedestrians' personal devices into the communication system [17, 19]. Yet, the practicality and effectiveness of these proposed solutions in multi-pedestrian settings have not been empirically evaluated or compared with each other [66]. Furthermore, while considerable research efforts have been directed towards understanding the influence of group dynamics in the context of AVs and eHMIs [8, 10, 11, 32, 44], the potential peer influences among co-located individuals remain underexplored. This knowledge gap is further exacerbated by the prevalent use of computer-operated virtual agents in studies, a method that may not fully capture the dynamic and nuanced aspects of real human interactions. To address the gaps identified around eHMI communication clarity and the dynamics among co-located pedestrians, this study poses the following research questions (RQ):

- **RQ1**: How do eHMIs on vehicles, infrastructure, and pedestrian devices target communications to intended recipients in multi-pedestrian settings?
- **RQ2**: How does the crossing behaviour of one pedestrian influence the behaviour of other co-located pedestrians in shared traffic environments?

We addressed these questions through a within-subjects study (N=25) comparing three different eHMIs, each representing a distinct communication locus. Our study utilised a multi-pedestrian virtual reality (VR) simulator, designed to accommodate two pedestrian participants simultaneously. VR simulations are common practice in AV external communication research as they provide a safer and more cost-effective alternative to real-world testing [50, 59, 65]. Employing a multi-pedestrian simulator allows for a better understanding of pedestrian interactions, thus adding realism to our research design.

This study contributes novel insights into AV communications by being the first to evaluate eHMI concepts across different loci—vehicles, infrastructure, and pedestrians—in dynamic multi-pedestrian environments. Our findings highlight communication challenges in the vehicle-based eHMI design, while infrastructure and pedestrian-oriented eHMIs showed better and more consistent performance across key metrics. We observed diverse pedestrian behaviours, ranging from those maintaining autonomy to those influenced by nearby individuals. Clearer AV instructions were found to reduce this influence on decision-making.

2 RELATED WORK

2.1 External Human–Machine Interfaces

2.1.1 Clarity of Recipients. The transition towards autonomous driving has brought the communication between AVs and pedestrians into sharp focus, with eHMIs identified as a promising tool to facilitate this interaction [54]. Our research investigates the efficacy of eHMIs in multi-pedestrian scenarios. In the taxonomy paper by Dey et al. [17], various dimensions were defined to categorise 70 different eHMI concepts. A particular dimension was the communication resolution, which refers to 'whether the eHMI concept enables the road user to identify with a certain level of detail or clarity for whom the message is meant'. Remarkably, out of 70 concepts analysed, only nine (representing 13%) offered high-resolution communication.

Several empirical studies have since delved deeper into this potential issue. For instance, a video study conducted by Wilbrink et al. [71], revealed a decreased willingness to cross among participants when virtual pedestrians were present on either side, contrasting with one-on-one encounter scenarios. Similarly, in a VR experiment by Dietrich et al. [20], a participant was asked to cross a road while an AV signaled it was safe to cross for the virtual pedestrian from the opposite side. The study's findings revealed that when the vehicle used undirected light signals, pedestrians are likely to interpret the communication as applicable to themselves. Consequently, the researchers recommend utilising directed signals or omitting signals altogether, which would allow pedestrians to rely on the vehicle's movement to make decisions.

Further investigation by Dey et al. [19] examined four distinct eHMI concepts in scenarios involving two pedestrians in VR. Participants stood on the pavement along with a virtual pedestrian who was positioned on the same road side, but 10 m apart. The study's findings suggest that eHMIs generally bolster the willingness of pedestrians to cross the road, even in situations when the communication message was not specifically intended for them. Using a WebGL application, Colley et al. [12] conducted a more comprehensive comparison of nine eHMI concepts in scenarios involving four-lane streets. The experimental conditions included scenarios where participants crossed the road either alone or with the presence of two virtual pedestrians, who were positioned 5 m to the left and right of the participant. The AV could yield for any of the three pedestrians. The study revealed that participants were less willing to cross in scenarios where AVs drove past and stopped for the pedestrian on the right. In these situations, participants felt significantly less satisfied and safe regarding their interaction with the AVs.

These studies collectively underscore the necessity for precise, high-resolution eHMI communications in multi-pedestrian settings. Promising eHMI concepts such as the Street Projection—which illustrates the AV's exact stopping point—have been identified as particularly effective [12, 19]. Our study aims to extend these findings by investigating a more comprehensive solution space that encompasses all communication loci.

2.1.2 Communication Locus. The *locus*, or physical location from which the communication originates, is a crucial dimension within the eHMI design space (alongside *message type* and *modality*) [13]. On-vehicle eHMIs, integrated into the AV's exterior, are a common solution; a 2020 taxonomy study found 89% of eHMI concepts utilise this design approach [17]. However, these eHMIs are generally designed to broadcast messages to a broad audience rather than individual recipients [17]. As a result, they may not always address the specific needs of each user, leading to possible confusion. Even when these eHMIs convey situational awareness by representing nearby pedestrians as simple forms, such as dots [69] and strips [52], their efficacy is bounded by their form factor [17]. Meanwhile, scalable solutions like the Street Projection concepts [12, 19], where an AV projects a zebra crossing onto the road, are noted for their scalability and positive user experience. However, these projection-based eHMIs also face challenges, particularly environmental factors like lighting and road conditions, which can affect their overall effectiveness [12, 19, 51].

To address these limitations associated with on-vehicle eHMIs, researchers are exploring the potential of alternative communication loci, including infrastructure and pedestrians' personal devices [17, 19, 27, 42, 62]. While promising, both approaches come with their own drawbacks [66]. Infrastructure-based eHMIs may involve high installation and maintenance costs as well as challenges integrating seamlessly into urban environments. Conversely, pedestrian devices raise concerns about accessibility, battery life, and privacy. The user experience implications of these alternative approaches are still largely unexplored, underscoring the need for further research to fully understand their potential and challenges. While studies exist that compare eHMIs across different loci (for example, between vehicles and infrastructure [27, 42], vehicles and pedestrian devices [53], and infrastructure and pedestrian devices [67]), our study offers a unique contribution by being the first to comprehensively compare all three loci in a multi-pedestrian context.

2.2 Multi-Pedestrian Simulators

Research on multi-pedestrian scenarios often employs virtual pedestrians in various prototype formats, such as VR [10, 19, 40, 44], computer videos [71], and web-browser games [12]. Researchers can introduce controlled variables into these scenarios by programming computer-managed pedestrians to cross the street at different times, disregard AV signals, or stand still [10, 19, 40, 44]. This approach allows researchers to observe and analyse how a human participant reacts to a range of other pedestrian behaviours.

However, using solely virtual pedestrians may not adequately represent the intricacies of real-world human interactions. In contrast, coupled or distributed simulators [4, 33, 34, 57, 75] offer a more nuanced approach by facilitating the study of interactions between pedestrians and occupants of both manually driven and AVs. These simulators provide a more realistic portrayal of the behaviours of different road users, although their application in multi-human pedestrian scenarios has not been observed.

The introduction of a multi-pedestrian simulator could offer significant advantages. Primarily, it enhances the social presence in VR scenarios, which refers to the experience of sharing a space with another individual perceived as real [46]. While pre-programmed computer agents could exert social influence that may approximate real-world interactions [30, 36], human-controlled avatars elevate this aspect by offering a more dynamic and realistic interaction experience in VR. For instance, human participants might show natural hesitation, choose unexpected paths, or interact with the environment in unpredictable ways.

3 DESIGN CONCEPTS

This section introduces three design concepts selected for evaluation within a multi-pedestrian VR simulation. Each concept represents a different communication locus—*Vehicle*, *Infrastructure*, and *Pedestrian* device—and has been identified as state-of-the-art in its respective category, based on prior research demonstrating its effectiveness.

3.1 Light Band + Street Projection eHMI (Vehicle)

For the *Vehicle* concept, we adapted the Light Band + Street Projection eHMI concept from a related study by Dey et al. [19]. This particular study evaluated four different versions of the light band concept in multi-pedestrian scenarios. Among these, the variant incorporating a projected crossing was found to be the most effective in conveying an AV's intention to yield. Its effectiveness was evident from pedestrians' lower willingness to cross when the AV stopped for others. Moreover, it notably improved pedestrians' experiences and their ability to discern whether the AV was yielding specifically to them or to another pedestrian. Our adapted design features an LED light band on the AV's bumper, displaying three distinct states: a solid light for cruising mode, an inward sweeping animation to signal yielding to pedestrians, and a projected zebra crossing on the road when at a complete stop (see Figure 2B and C). The light band is turquoise, in line with colour recommendations from existing literature [3, 18, 21], while the zebra crossing is green, consistent with the colour schemes of the *Infrastructure* and *Pedestrian* concepts. Another deviation from the original concept involves omitting the projection of arrows that indicate ahead the AV's stopping point, a change necessitated by the challenges of dynamic long-distance projection.

Previous studies suggest that eHMIs may inadvertently increase pedestrian willingness to cross, even when the AV is yielding for someone else [19]. To mitigate this, we integrated a late activation in this eHMI concept. This delay necessitates pedestrians to closely observe the AV's movements. Specifically, the AV starts braking at a distance of 40 m but only signals its intent to yield when it is 20 m from the pedestrian.

3.2 Smart Curbs (Infrastructure)

Several eHMI concepts, including projection-based drones [25], adaptive road surfaces [68], and smart curbstones [27], incorporate infrastructure elements. We chose the Smart Curbs concept for the *Infrastructure* category for several reasons. First, it integrates seamlessly into existing urban environments, minimally impacting city planning and construction. Second, it employs a system similar to in-ground LED lighting to enhance road safety, particularly for pedestrians absorbed in their mobile devices while crossing streets[1]. Crucially, it has undergone formal evaluation [27] and shows promise in addressing the scalability challenge in AV–pedestrian communication by localising information to pedestrians' current positions.

The Smart Curbs concept uses a simple two-colour scheme (red and green) to indicate to pedestrians the safety of crossing at a specific location. Unlike the original design where the curbstones spanned the entire roadside, our adaptation involves lighting up only in areas with pedestrians. This change aims for economic efficiency and reduces the potential visual overload of continuous lighting. The change from red to green also becomes a direct switch rather than a gradual transition that follows vehicle movement. This simplification aims to make the system more predictable for pedestrians (see Figure 2D and E).

3.3 Smartglasses (Pedestrian)

While various pedestrian devices, such as smartphones [28, 29, 45, 47, 73] and wearable accessories [39], can be utilised to improve pedestrian safety, wearable Augmented Reality (AR) technology, particularly smartglasses, is gaining prominence due to its seamless integration with the user's visual field, supporting retention of situational awareness [2]. As a result, wearable AR has been extensively discussed and investigated in the recent literature of automated driving, with a focus on both contexts: in-vehicle usage [55, 60] and usage by external road users [24, 53, 61, 62, 67].

Several research studies have explored the use of augmented zebra crossings as a key feature in smartglasses to enhance pedestrian safety [61, 67]. These studies highlight that zebra crossings are universally recognised road markings, making these crossings ideal for conveying safety messages to pedestrians of diverse backgrounds. Therefore, this visual element of an augmented zebra crossing forms an essential part of the *Pedestrian* AR concept. Additionally, our *Pedestrian* AR concept also features a head-locked warning known as the 'Nudge Head-up Display (HUD)'. This design has been demonstrated to be the most intuitive among nine novel AR interfaces for pedestrian–vehicle interaction, as evaluated in an online study with 992 respondents [61]. This interface projects a 'stop' hand symbol and a warning message 'Danger! Vehicle is approaching' directly into the user's field of

[1] https://www.koreaherald.com/view.php?ud=20220111000281

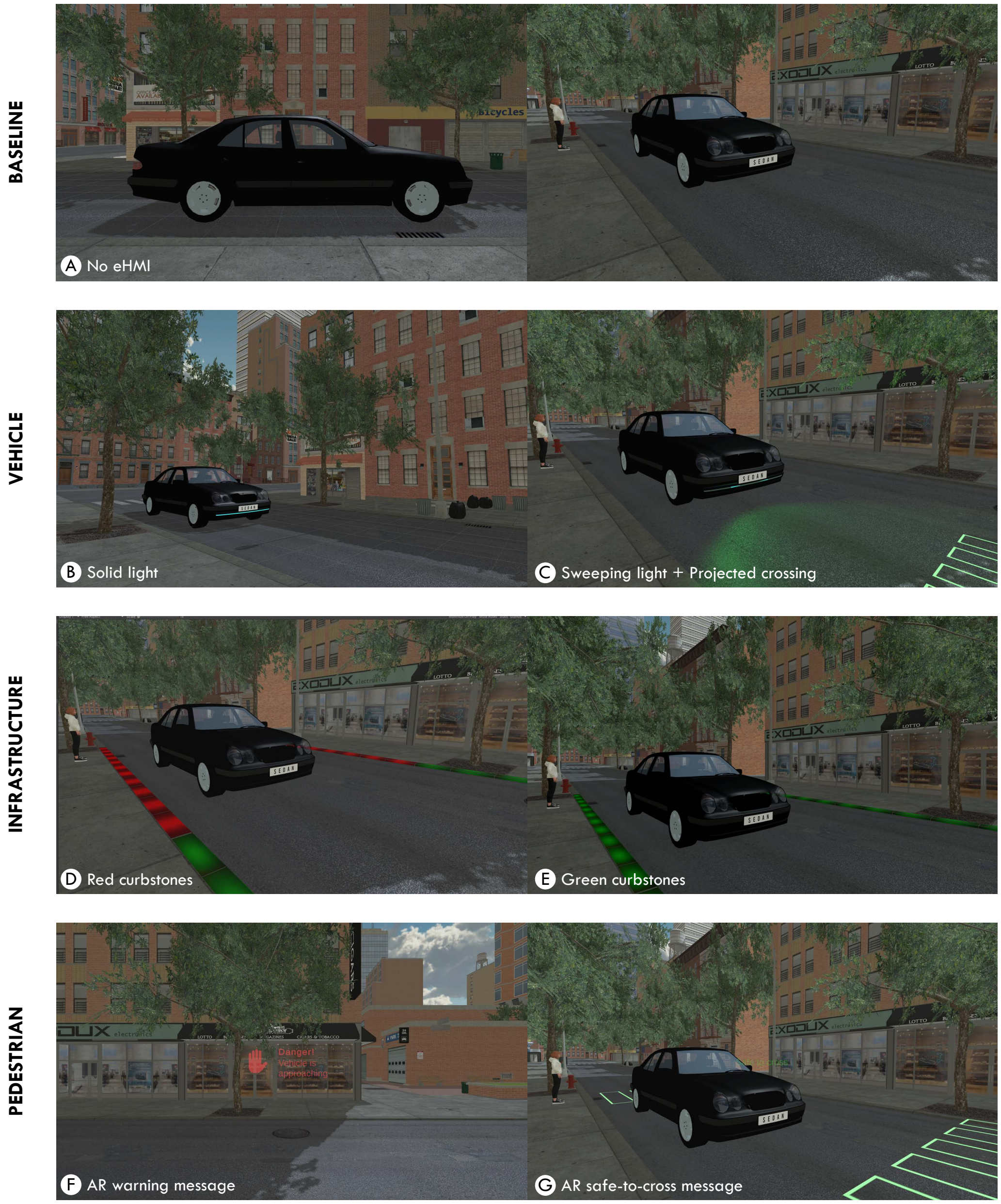

Fig. 2. Interface conditions: Baseline, Vehicle, Infrastructure, and Pedestrian (displayed from top to bottom). In the Baseline condition, the AV has no eHMI (A). In the Vehicle condition, the AV features a solid light while cruising (B), transitions to a sweeping light upon braking, and projects a zebra crossing when fully stopped (C). In the Infrastructure condition, the curbstone lights up red to indicate it is not safe to cross at a given position (D) and green when it is safe (E). In the Pedestrian condition, pedestrians are alerted with a warning message as the AV approaches (F), a safe-to-cross message during AV braking, complemented by an augmented zebra crossing when the AV comes to a full stop (G).

view. The message updates to 'Safe to cross' when it is safe to do so. Since our study involves two pedestrian participants, we implemented a shared AR experience[2] in our design. In this setup, each participant receives an individual warning about approaching vehicles, while the augmented zebra crossing will be a shared visual element, visible to both participants (see Figure 2F and G). The shared AR experience allows both pedestrians to be aware of how the AV responds to each of them.

4 EVALUATION STUDY

4.1 Study Design

The virtual setting represents a two-way urban street devoid of traffic lights or marked crossings, creating an ambiguous scenario where rights of way are not clearly defined. Participants were spaced 10 m apart on the pavement, as per Dey et al. [19], to represent a situation where pedestrians are co-located but not in a group. Their task involved carefully observing the approaching AVs and the surrounding environment, then deciding whether to cross the road. Importantly, participants were informed that not crossing was an option and would not affect the study's outcome.

We implemented a within-subject design for our study. Each participant experiences four different conditions: three involving varying eHMI concepts (*Vehicle*, *Infrastructure*, and *Pedestrian*) and one baseline scenario without an eHMI. These conditions are organised into four blocks, with each block comprising five different scenarios. To counteract potential order and carryover effects, we use a balanced Latin square method [6] for randomising the sequence of eHMI conditions and the scenarios within each block.

Each scenario block introduces different vehicle behaviours: a) the AV yielding to the participant, b) the AV yielding to the other pedestrian, and c) the AV not yielding. Additionally, participants alternate between two positions in these scenarios. In the first position, the participant is nearer to the oncoming AV and faces a higher risk of collision due to possible misinterpretations of the AV's signals. In the second position, the participant is further from the AV and has a view of the other pedestrian. These roles allow participants to experience varied risk and influence factors. Please refer to Figure 3 for an illustration of the scenarios.

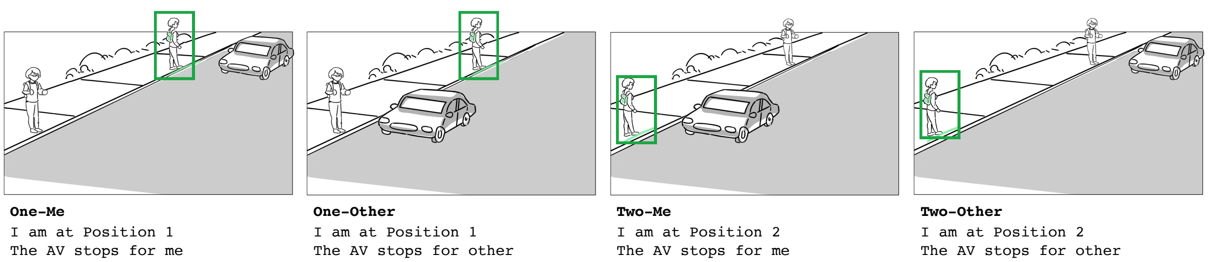

Fig. 3. Four scenarios illustrated from the perspective of a participant (excluding the scenario with a non-stopping AV).

4.2 Apparatus

4.2.1 Hardware. Participants were provided with a Meta Quest 2 headset, which offers an untethered VR experience. The device has a resolution of 1832 x 1920 pixels for each eye and a diagonal field of view estimated between 90 and 100 degrees. For hand-tracking purposes, participants were asked to hold two Touch controllers throughout the study.

[2]A shared AR experience refers to a scenario where multiple users can see and interact with the same AR content simultaneously.

4.2.2 Virtual environments. For the study, we designed three distinct virtual environments (VEs): a lobby area, a familiarisation VE and a test environment. In the lobby area, participants could select an avatar and connect to the system (see Figure 4 right). Both the familiarisation environment and the test environment were designed to mimic real-world urban environments. However, an AI-driven vehicle was introduced only in the test environment. The vehicle is spawned at a position outside the participant's line of sight, approaching from the left at 50 km/h and decelerates at 2.4 m/s^2, covering a 40-m distance in approximately 5.3 seconds. The sound of the vehicle's engine is perceived spatially relative to the participant's position.

4.2.3 Networking. The multi-pedestrian VR simulator was developed using the Unity[3] game engine, incorporating 3D assets from the Unity Asset Store. Multi-user networking was implemented via the Photon Unity Networking (PUN)[4] framework, specifically PUN 2. To mitigate latency issues, we employed strategies such as a high-speed internet connection and VR optimisation techniques, which reduced graphic processing demands and enhanced user synchronisation.

4.2.4 User Role. The simulator features three roles: Player 1, Player 2, and Observer. The Observer is rendered invisible and maintains a preset viewing angle, allowing them to monitor both VR participants. This specific angle is captured for subsequent analysis (see Figure 4 left). As the Master Client[5] in our network, the Observer also manages scenario progression. The role was operated using a 2020 iMac.

4.2.5 User Avatar. Participants were represented by full-body avatars from Ready Player Me[6], as opposed to avatars lacking lower body representations. While in social VR applications (e.g., Horizon Worlds [49] and Microsoft Mesh [15]), legless avatars may suffice, full-body representations are critical in simulating pedestrian behaviours. The inclusion of legs in avatars is instrumental for conveying movement, direction, and speed, essential for realistic pedestrian dynamics.

Initially, we set up a process allowing participants to upload their photos and personalise their avatars, motivated by study findings where individualised avatars could support a greater virtual body ownership [31, 70] and boost users' sense of embodiment and social presence [35]. However, during our pilot tests involving four participants, it became evident that personalising avatars for individual participants was less impactful than anticipated, primarily because participants could not see themselves during crossing scenarios. Moreover, with participants having no prior relationship with each other, the identity that is typically important in a social setting played a less pronounced role. For these reasons, we opted for pre-existing and non-personalised avatars to simplify preparation tasks for study participants and reduce technical issues that might occur during the process of loading individualised avatars. We used one female and one male avatar (see Figure 4 left), usually assigned corresponding to the participant's gender. However, in same-gender pairings, one participant was allocated an avatar of the opposite gender for differentiation purposes. Findings from the pilot tests, as well as in instances of same-gender pairing, confirmed no impact of mismatched gender avatars.

The Meta Quest 2 is capable of tracking users' head and hand movements. Utilising inverse kinematics [1], the simulation estimates the position of elbows, torso, and legs, enabling users to exhibit lifelike movements within virtual scenarios. Although this method does not achieve the precision of full-body motion suits [37], it offers a substantial reduction in both setup cost and complexity. Facial expressions of the avatars remain neutral and unchanged; however, the eye animation (e.g., blinks) was utilised to make them look alive.

4.2.6 Concept Implementation. The *Vehicle* concept utilises a bespoke 3D model for the light band, crafted to resemble real-world implementations [43]. For the *Infrastructure* concept, an LED-like emissive material was

[3] https://unity.com/
[4] https://www.photonengine.com/pun/
[5] In networking scenarios, the Master Client typically has elevated permissions.
[6] https://readyplayer.me/

Fig. 4. (Left) Two positions in each scenario: the participant standing at Position 1 is closer to the approaching AV; the participant standing at Position 2 is further away from the AV. (Right) In the lobby area, participants could select their avatar and, by moving their bodies, control and observe its movements in a mirror.

used to mimic LED-embedded curbstones. For the *Pedestrian* concept, it was crucial to ensure that AR HUD messages are non-obstructive and allow participants to remain attentive to the 'real' world. To this end, the interface was made semi-transparent, with alpha (opacity) values set at 120. The messages are also designed to disappear when the participant starts crossing, reducing potential distractions. Regarding the zebra crossing, both *Vehicle* and *Pedestrian* concepts employ emissive materials for a glowing effect. In the *Vehicle* concept, the crossing is confined to the area in front of the vehicle, and is complemented by a simulated laser projection from the vehicle's bumper.

4.3 Participants

Our study initially involved 26 participants, recruited via social media and word-of-mouth. Eligibility criteria included being over 18, English fluency, normal or corrected vision, and no mobility impairments. Sessions were conducted in pairs. Unfortunately, one participant withdrew post-experiment, precluding their involvement in the subsequent interview and necessitating the cancellation of that session's group discussion. Nevertheless, the data from their partner was retained, resulting in a final participant count of 25 (see Table 1). Ethical approval was granted by the university's human research ethics committee, and participants received $20 as compensation.

To explore the influence of co-located pedestrians (as per RQ2), we strategically paired participants who were strangers. Preliminary pilot tests showed that pre-existing relationships influenced interaction and behaviour in both virtual and real-world contexts. Pairing strangers mitigated these biases, fostering independent decision-making and cautious interaction due to unfamiliarity with fellow participants.

4.4 Study Procedure

Upon registration for the study, participants were asked to specify their availability and confirm their allocated timeslot. Additionally, they needed to verify that they had no previous acquaintance with the other individual assigned to the same session. A day prior to the session, they were provided with prior information on the design concepts, each illustrated through a brief video and an accompanying short description (see Appendix A).

Each study session was conducted by two researchers, with each researcher assigned to one participant. To minimise direct real-life interactions, participants were separated upon arrival and communicated exclusively with their assigned researcher until they reached the group discussion segment of the interview. The sessions were conducted in a spacious studio, divided into two areas of 8 x 5 m each (see Figure 5 for layout). Participants were

Table 1. Demographics and prior experience of participants.

Category	Details
Gender (m/f/unspecified)	12/12/1
Ages	
18—24	6
25—34	17
35—44	1
55—64	1
Vision Correction	16
VR Experience	23
AV Experience	
Read about AVs	12
Interacted with AVs	5
None	5
Conducted research	3

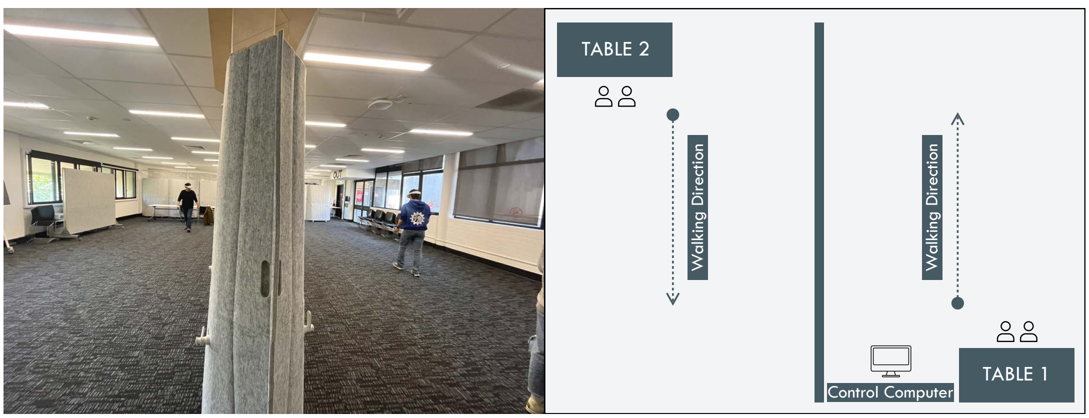

Fig. 5. Room setup: A physical divider splits the studio in half, separating the participant-researcher pairs. There is Table 1 for the first pair and Table 2 for the second. The control computer is located at Table 1.

briefed on the study procedures before signing a form to confirm their voluntary participation and completing a demographic questionnaire. This was followed by a familiarisation phase, where participants were able to see each other and practise crossing streets in VR. This initial step aimed to mitigate any discomfort induced by VR and build confidence in participants for the experiment.

To ensure that both participants began each scenario simultaneously, a researcher managed the transitions between scenarios using a computer. Participants received an in-app notification stating, *'Moving to the next scenario'*, before being transitioned to the subsequent one. Following every block of five scenarios, participants were prompted to discuss their recollections of the scenarios or explain any specific behaviours noted by the researcher. They then filled out a series of questionnaires. At the conclusion of the study, they completed a presence-related questionnaire and took part in a semi-structured interview (see Figure 6).

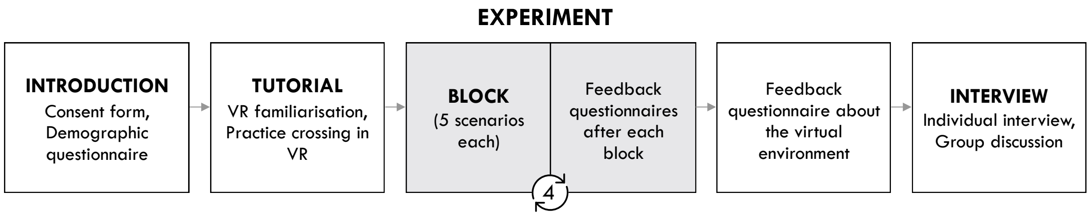

Fig. 6. Study procedure.

4.5 Data Collection

Our data collection strategy encompasses both quantitative and qualitative assessments to understand pedestrian crossing performance, specific intent understanding, and interaction experiences.

Crossing performance: HMD-logged data were collected to determine crossing initiation time (CIT)—the time measured in seconds between AV deceleration and the participant's road entry [63, 74]. A lower CIT reflects an efficient decision-making process [63], to which communication clarity, among other factors, contributes. Instances of collisions were also recorded using HMD-logged data. Non-crossing decisions were identified through a combination of the researcher's observation and the participant's verbal confirmation. Additionally, we recorded the VR session from a third-person perspective, enabling the revisiting of participant interactions with the AV.

Feelings of being addressed: Measured by a 5-point Likert scale (1–very weak, 5–very strong), adapted from [71]. The question was phrased as 'How would you rate your experience of being personally addressed, meaning the AV's communication is directed specifically towards you and not another person?'

Workload: Measured by the NASA Task Load Index (TLX) scales [23] in six dimensions: mental, physical, temporal demand, performance, effort, and frustration. Each of these dimensions is rated on a scale from 0 (very low) to 100 (very high), used to compute an overall workload score.

User experience: Measured by the short version of the User Experience Questionnaire (UEQ-S) [41] to assess participants' user experience with the AV. The scale comprises eight pairs of opposite adjectives: four pairs representing Pragmatic Qualities and four representing Hedonic Qualities.

Trust: Measured by the Trust In Automation questionnaire by Körber [38]. Our primary focus was on the overall trust scale, as well as two specific subscales: Reliability/Competence and Understandability/Predictability.

Motion sickness and presence: The Misery Scale [5] was used to monitor simulator sickness, suspending the study if ratings exceeded three. The Multimodal Presence Scale (MPS) [46], which measures physical, social, and self-presence, was employed to evaluate social presence in a multi-pedestrian simulator and the impact of a full-body avatar on self-presence.

Semi-structured interviews: At the end of the study, each participant was invited to rank the conditions (1–most preferred, 4–least preferred), and elaborate on their preferences. Our aim was to gain a deeper understanding of how they interpreted the signals and perceived the behaviours of other pedestrians in response to these designs. Subsequently, we facilitated group discussions to uncover common experiences, divergences in perceptions, and peer influences.

4.6 Data Analysis

4.6.1 Quantitative Analysis.
We used IBM SPSS Statistics Version 29.0.1.0 for all analyses.

Internal reliability: Cronbach's alpha was calculated to assess the consistency of the items within the questionnaires. High reliability indicates that the items measure the same underlying concept, which is essential for the validity of each scale. The NASA TLX showed good reliability with $\alpha = .890$. For the UEQ-S, both the

Pragmatic Qualities subscale and the Hedonic Qualities subscale demonstrated excellent reliability, with values of $\alpha = .913$ and $\alpha = .938$, respectively. Trust in Automation's overall score was excellent at $\alpha = .900$. The Reliability/Competence subscale showed good reliability at $\alpha = .835$, while the Understandability/Predictability subscale was acceptable with $\alpha = .773$. Lastly, the MPS scale's overall score was excellent at $\alpha = .908$. The subscales for MPS registered good reliabilities: Physical Presence at $\alpha = .827$, Social Presence at $\alpha = .811$, and Self Presence at $\alpha = .906$.

Descriptive analysis: To obtain a general understanding of the data and support the selection of appropriate inferential statistical tests, we conducted a descriptive analysis on the questionnaire data. This involved calculating the mean, median, and standard deviation for each scale and visually representing the distribution of scores to identify outliers. The normality of the data distribution for each scale was assessed using the Shapiro-Wilk test.

Statistical tests: Given the non-normal distribution of the questionnaire data, we employed the non-parametric Friedman test to identify statistically significant differences among conditions. For analysing crossing initiation time (measured per scenario), we applied the Linear Mixed Model (LMM) to account for the hierarchical structure of our study design—where scenarios are nested within each condition—and to address the repeated measures aspect. Within the LMM, conditions and scenarios were set as fixed effects, and participants were treated as a random effect to accommodate inter-individual differences. Of note, we excluded data from scenarios in which the AV did not stop. These scenarios were introduced solely to make the AV's behaviour less predictable [19, 65]. In total, data from 400 crossing trials (4 conditions x 4 scenarios x 25 participants) were analysed.

If significant results were found in either the Friedman test or the LMM analysis, we proceeded with post-hoc pairwise comparisons. To control for the increased risk of Type I errors due to multiple comparisons, we applied the Bonferroni adjustment method to these post-hoc tests. The adjusted p-values from these pairwise comparisons are reported and assessed for significance at $p < .05$.

4.6.2 Qualitative Analysis. Individual interviews and group discussions were transcribed using an AI transcription service[7]. The researchers who conducted the interviews reviewed and corrected the transcriptions. We applied thematic analysis to interpret the data [7] and utilised Miro[8], a visual collaboration platform, to support the analysis.

For individual interviews, the respective researchers who conducted them coded their segments independently. For group discussions, both researchers independently coded the entire data set. The process began with examining 25% of the data, after which they compared identified codes and themes and discuss differences. Once aligned, they moved on to analyse the rest. This concluded with a final round of discussions to finalise the themes.

5 RESULTS

5.1 Crossing Performance

5.1.1 Crossing Decisions. Out of 400 trials, there were 8 instances where participants decided not to cross. In the following, we present scenarios featuring these non-crossing decisions, along with the reasons behind them.

- P17: (Baseline - 3 times) Not crossing in front of a stopped AV without eHMI, mentioning that *'My instinct tells me that the vehicle might have stopped for me. However, there's no clear way to verify that.'* P17, however, did cross at the back of a stopped vehicle.
- P21: (Baseline - 2 times) Not crossing in front of a stopped AV without eHMI; in both scenarios, the AV had stopped for him. P21 stated, *'I was concerned about the possibility of the car suddenly starting up while I was crossing. Without any clear signals from the car indicating it's safe, I'd rather not risk it.'*

[7]https://otter.ai/
[8]https://miro.com/

- P19: (Vehicle - 2 times) Not crossing in scenarios where the AV stopped for the other pedestrian. P19 explained this decision as, *'When the car went past me and projected for that pedestrian, it felt very targeted at that specific person. I didn't think they could see me because their attention was directed that way'.*
- P21: (Vehicle - 1 times) Not crossing in the scenario where the AV stopped for the closer pedestrian. P21 mentioned he prioritised his safety, *'It didn't seem to stop for me specifically. I felt it might be dangerous if I tried to cross then'.*

5.1.2 Collision. There was one collision recorded in the Baseline condition, in the scenario where the AV did not stop for any pedestrians. The AV was travelling at a speed of 50 km/h, and P10 was the furthest pedestrian. P10 mentioned she would never do it in real life but thought the situation involved an AV and therefore wanted to test it: *'I think it's not smart because when I want to cross the road, it hits me'.*

5.1.3 Crossing Initiation Time. Out of 400 trials, there were 2 trials with negative CITs (i.e., participants began crossing before a vehicle started yielding). The LMM analysis demonstrated a significant effect of condition on CIT, F(3, 95.241) = 24.711, p < .001. Subsequent pairwise comparisons revealed significant differences in CIT between conditions. Specifically, the *Baseline* condition led to a significantly longer CIT compared to the *Infrastructure* condition (Mdiff = 3.500, SE = 0.462 , p < .001), and the *Pedestrian* condition (Mdiff = 2.530, SE = 0.579, p < .001). Furthermore, the *Infrastructure* condition was associated with a significantly reduced CIT compared to the *Vehicle* condition (Mdiff = 2.205, SE = 0.401, p < .001). There is also a significant interaction effect of Condition and Scenario, F(12, 39.063) = 12.367, p < .001 (see Table 2).

Table 2. Crossing initiation time: Estimated marginal mean value (**M**), standard deviation (SD), and pairwise comparisons across conditions.

	One-Me (**M**/SD)	**One-Other** (**M**/SD)	**Two-Me** (**M**/SD)	**Two-Other** (**M**/SD)
Baseline	**10.95**/1.13	**11.84**/1.18	**8.21**/0.53	**8.88**/0.43
Vehicle (V)	**8.00**/0.37	**10.01**/1.20	**8.24**/0.44	**8.45**/0.71
Infrastructure (I)	**6.20**/0.54	**5.68**/0.51	**5.64**/0.47	**8.36**/0.32
Pedestrian (P)	**6.28**/0.45	**7.28**/1.50	**7.16**/0.51	**9.04**/0.32
	Baseline > I (p = .002) Baseline > P (p = .002) V > I (p = .012) V > P (p = .003)	Baseline > I (p <. 001) V > I (p = .010)	Baseline > I (p <. 001) V > I (p <. 001)	P > I (p <. 046)

5.2 Feeling Addressed

The Friedman test indicated a significant difference in scores among the conditions, ($\chi^2(3) = 21.681, p < .001$). Pairwise comparisons revealed that participants felt significantly less addressed in the *Baseline* condition than in the *Infrastructure* ($p = .037$), *Vehicle* ($p = .022$), and *Pedestrian* ($p = .001$) conditions (see Figure 7).

5.3 Workload

A significant difference in the median workload scores across the four conditions *Baseline*, *Vehicle*, *Infrastructure*, and *Pedestrian* was observed, with $\chi^2(3) = 16.671$, $p < .001$. Further pairwise comparisons revealed that both the *Pedestrian* and *Infrastructure* concepts had significantly different median scores compared to the *Baseline*, with significance levels of $p = .002$ and $p = .003$ respectively (see Figure 7).

5.4 User Experience

The Friedman test revealed statistically significant differences in distributions across the conditions for both Pragmatic ($\chi^2(3) = 25.402$, $p < .001$) and Hedonic qualities ($\chi^2(3) = 36.590$, $p < .001$). Pairwise comparisons underscored that these differences were primarily between the *Baseline* conditions and the other respective conditions (*Vehicle*, *Infrastructure*, *Pedestrian*) (see Figure 7).

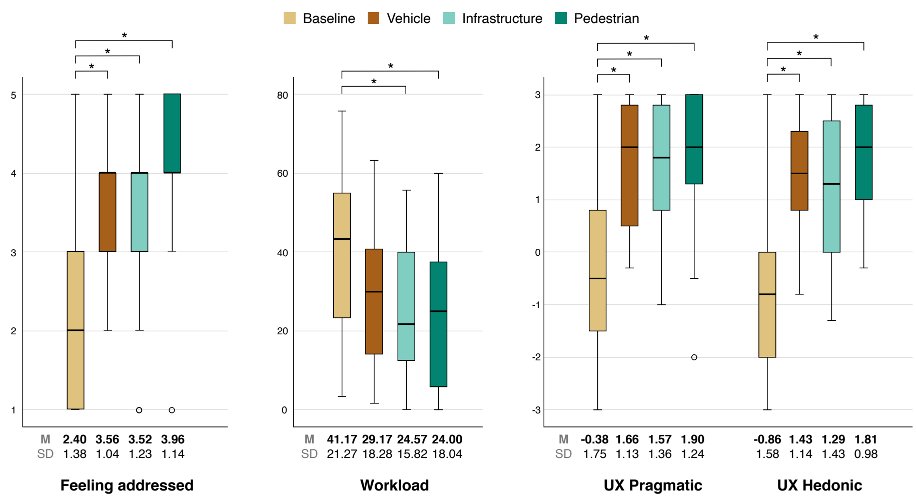

Fig. 7. Box plots of Feeling Addressed scores (left), Workload scores (middle), and User Experience (right) across different interface conditions. An asterisk (*) indicates $p < .05$.

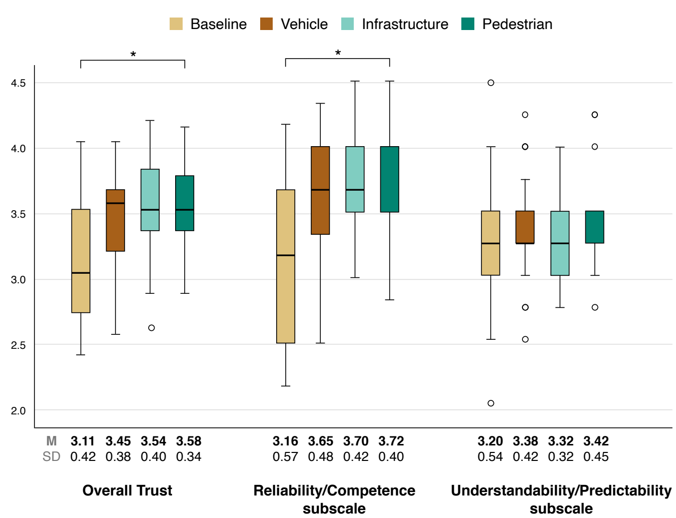

Fig. 8. Box plots of Trust and Trust subscales across different interface conditions. An asterisk (*) indicates $p < .05$.

5.5 Trust in Automation

The Friedman test has revealed a significant difference in the overall trust scores among the conditions: $\chi^2(3) = 10.371, p = .016$. Significant differences in trust scores were found between *Baseline* and *Pedestrian* ($p = .011$). Similarly, the test indicated significant differences across the conditions in the Reliability/Competence subscale ($\chi^2(3) = 16.703, p < .001$). In pairwise comparisons, only the scores for *Baseline* when compared to *Pedestrian* showed a statistically significant difference ($p = .001$). For the Understandability/Predictability subscale, the results indicated that there were no significant differences in the distributions of scores among different interface conditions (see Figure 8).

5.6 Preference Ranking

The *Baseline* condition is clearly the least preferred ($M = 3.60$). In contrast, the *Pedestrian* concept stands out as the most preferred ($M = 2.04$). Both the *Vehicle* concept and *Infrastructure* concept have similar moderate preferences, with mean rankings of 2.20 and 2.16, respectively.

The Friedman test revealed significant differences in participants' rankings across the conditions, ($\chi^2(3) = 24.408, p < .001$). Pairwise comparisons demonstrated that the *Baseline* was significantly less preferred than the *Vehicle* concept ($p = .000$), the *Infrastructure* concept ($p = .000$), and the *Pedestrian* concept ($p = .001$).

5.7 Presence

In the MPS questionnaire, participants consistently rated Physical Presence highest ($M = 3.94$, $SD = 0.66$). Social Presence had a slightly lower average ($M = 3.72$, $SD = 0.82$), while Self Presence showed the most variability and the lowest mean score ($M = 3.44$, $SD = 1.04$). This indicates a generally good sense of presence, with the most uniform experiences in Physical Presence and the most varied in Self Presence.

5.8 Qualitative Findings

In this section, we outline participants' experiences with various interface conditions. We also detail the influence of co-located pedestrians on crossing behaviours, their feedback about the AV behaviour, and the multi-pedestrian VR simulation.

5.8.1 Perception and Experience of Interface Conditions. Each interface concept will be discussed in terms of general experience themes, followed by themes associated with direct and targeted communication.

Baseline. In the *Baseline* condition, participants frequently highlighted ambiguity and uncertainty about the AV's intentions (n=6), leading to an increased cognitive load (n=6). Safety concerns were evident, with feelings of being hurried (n=2) or undetected (n=3). This was compounded by a clear lack of trust in the AV's actions (n=3). Despite these concerns, there was also an appreciation for its realism and familiarity (n=8). Participants often commented on the AV's likeness to a regular car, with one noting, *'It seems like a normal vehicle with a driver inside,'* (P8) and another drawing parallels to *'walking across the street every day'* (P12). Additionally, this condition instilled a sense of confidence and autonomy in some participants (n=6), stating sentiments like *'I really felt more confident without the augmentation'* (P19) and *'I can do whatever I want'* (P12).

Participants notably relied on the AV's movement as a primary cue to ascertain if the AV had stopped for them (n=6). This behaviour brought about moments of ambiguity and hesitancy. P1 remarked, *'When it finally came to a stop, it was clear. But leading up to that moment, it wasn't evident'*. It underscores that, in the absence of explicit interface cues, the full stop of the AV served as the most definitive indication of its intention toward the pedestrian.

Vehicle. The presence of eHMI on the AV indicated awareness of pedestrians, and left an impression of the vehicle's *'intelligence'* (P16). This condition was also associated with feeling that interactions were less mentally demanding (n=3), and participants found it to be safer (n=6). The direct cues from the vehicle were perceived as reliable due to their source (n=4). As P22 noted, *'I feel like it all comes from the car, it doesn't have to communicate with the public infrastructure or the smartglasses I'm wearing. So it might have less chance to get errors'*. In addition, it made the pedestrians more observant (n=2). P8 articulated, *'I turned my head around to see to find the car'*.

In terms of targeted communication, participants often associated the vehicle's projected crossing as being intended for a particular individual (n=7). As P20 observed, *'It's like a pathway created for me. I feel it is communicating with me.'* However, this direct communication also introduced ambiguity for nearby pedestrians who were not the primary target of the AV's signals (n=5). It unintentionally caused other pedestrians to question if they were noticed by the AV. P17 voiced these concerns, stating, *'All I could tell is that it had detected one person and it put a pedestrian crossing in front of them, but there's no way for me to verify whether the vehicle knows that I'm also further down and also crossing'*. We noted a few instances where participants misinterpreted the light band (n=2). P5 mentioned, *'It was unsettling when the car signaled it would stop for a pedestrian but continued past me. I almost walked into it, thinking it would stop for me'*.

Infrastructure. Many participants found the curbstone design to be interesting and novel (n=3), and it was generally regarded as easy to understand (n=2), making it accessible to a wide range of users. The design showed similarities to existing infrastructure components, such as traditional crosswalks and traffic lights (n=7).

Perceived connection between infrastructure and vehicles emerged as a recurrent theme where participants voiced varied perceptions (n=10). P5 voiced uncertainty, stating, *'I'm not sure if it's the vehicle triggering it or something else.'* P16's observation, *'the capstone gives the impression it's doing its own job, separate from the autonomous vehicle's function'*, reinforces this notion. P19 speculated on the technology involved, *'I assume there's machine vision or sensors detecting the vehicle's speed and direction relative to a pedestrian.'* There was also a sense that, as an infrastructure-based concept, the design might hold more authority (n=3). However, this perception of the infrastructure acting as a third-party or external system led participants to express the need to distribute their attention between multiple cues (n=5). P5 encapsulated this sentiment by noting, *'I found myself needing to actively check for oncoming vehicles. I have to check the status of the curbstone and then determine if it was safe based on that and other cues.'* Skepticism also arose about its city-wide implementation.

Despite the curbstone's potential to communicate unambiguously with multiple pedestrians [27], participants encountered confusion due to the colour split between red and green (n=15), which is intended to signal which pedestrians the AV is yielding for. P5 noted, *'this situation was especially odd [...] it felt as though I was in a large pedestrian crossing, but then the lights on the pedestrian crossing malfunctioned'*. P2 added, *'the curbstone makes me feel like it's an entire area just for pedestrians, but it's actually divided into two parts.'*

Pedestrian. Many participants appreciated the clarity and reduced cognitive workload the technology offered (n=10). P11 stated, *'this one is very clear. It just tells you when it's not safe and where is safe. It immediately projects the crosswalk.'* The glasses also seemed to enhance users' environmental awareness beyond traditional human capabilities (n=3). P17 noted, *'even if I wasn't directly observing the car, an alert indicating a "dangerous vehicle approaching" would appear,'* highlighting the glasses' potential to heighten vigilance. Moreover, trust in this technology emerged as a significant theme (n=5). P9 shared, *'Even if the car is still a distance away, I get a warning. So, when I see a green signal, I proceed without hesitation. It's a matter of trust in the system.'*

While the smartglasses were praised for its clear instructions, it also seemed to impact the pedestrians' sense of autonomy. Some participants felt that the technology might be somewhat redundant or even override their natural instincts and observations (n=4). P19 voiced a nuanced perspective: *'The smartglasses told me danger, there was traffic coming [...] I thought it was useful, but at the same time, I'm quite capable to do that myself.'* This sentiment was underscored by other participants who found themselves relying heavily on the glasses' guidance, occasionally to the detriment of their own judgment. P8 recalled a specific instance: *'there was one point when*

I tried to cross the street. The instructions from the glasses, not the real situation with the car, guided me. I didn't decide based on what I saw in reality; I just went by the instructions from the glasses.'

In this concept, participants also highlighted the connections between the technology, vehicles, and the overarching city infrastructure (n=4). P1 envisioned an integrated future, stating, *'if vehicles become smarter, they can communicate with glasses. The glasses could also determine where the vehicle was and indicate when it's safe to cross.'* P17 elaborated on this interconnectedness, emphasising a shift from a car-centric system to a broader smart city perspective: '*I was thinking more of a smart city where I have a heads-up display, it's this technology that I'm wearing, and perhaps the city itself is the operating system that all of these interactions are working on.* This perception of a comprehensive, interconnected system gave P17 a feeling of being *'reassured'* by the AR glasses, sensing that the entire environment was designed to ensure their safety.

The smartglasses displayed a remarkable ability to communicate distinctly with individual pedestrians (n=15). P14 stated, *'[the message] was popping up in my glasses and I knew it was to me.'* Beyond this personalised acknowledgement, participants also recognised the system's capability to identify and engage with multiple pedestrians concurrently via the shared AR elements (n=6). This broader awareness not only demonstrated technological proficiency but also played a role in shaping user trust. P9 reflected on this multi-user recognition, mentioning, *'I understand the car recognises she's in its blind spot, so I trust the car more.'*

5.8.2 Co-Located Pedestrian Influence on Pedestrian Behaviours. Co-located pedestrian influence varied among participants. For some, the presence and actions of other pedestrians had minimal to no impact on their decisions (n=9). P8 emphasised their autonomy in decision-making by saying, *'I make my own decisions. Not even slightly influenced.'* The presence of another individual in the VR environment initially piqued P7's interest, possibly because of the novelty of the experience. They remarked, *'Initially, I was curious about another person in the VR environment. But later on, not really.'* Moreover, the proximity of co-located pedestrians played a crucial role *'only if we were standing next to each other'*.

Meanwhile, many participants acknowledged the impact of co-located pedestrian influence (n=12). For some, the effect might be subtle and underlying. For instance, P4 conceded that, while their decisions were not driven by others' actions, there might have been a *'subconscious'* reduction in their mental workload. This influence was more overt for P17, who shared, *'Even just with one other person, there was a little bit of that dynamic, kind of like we're both coming to a decision about what is the appropriate level of caution.'* Interestingly, remarks regarding co-located pedestrian influence predominantly related to scenarios where the AV stopped for other pedestrians (n=11). This includes instances when participants were crossing behind the vehicle or when they were crossing at a considerable distance from the vehicle. They relied on their co-located pedestrian due to the uncertainty that arose when not being acknowledged by the AV. As a result, they felt safer initiating a crossing simultaneously with the pedestrian who was the target of the AV's attention.

The perception of the decision-making process varied based on the two crossing positions. The position closer to the vehicle was often seen as more challenging (n=7). This was attributed to the reduced time available to make decisions, combined with the need to make these decisions independently. On the other hand, the farther position, while still perceived as challenging, presented a different set of complexities (n=5). Here, pedestrians felt the need to consider a broader set of information in their decision-making process, encompassing not just the vehicle's actions but also those of co-located pedestrians.

The analysis revealed that the influence of co-located pedestrians varied based on the interface conditions (n=7). A contributing factor was the varying degrees of instructions provided by each interface, leading to uncertainties as participants attempted to interpret different situations. By design, some concepts, such as smartglasses, provided more instructions and earlier on, in contrast to others like the no-interface concept which delivered fewer, later-stage instructions. P5 and P6 mentioned they were more attentive in situations that did not involve AR. In contrast, P9 observed, *'it doesn't really matter for the curbstone and smartglass concept. Regardless* *of whether there's a pedestrian or not, I would just cross.'* However, not all participants felt the same way. Some, like P24, were skeptical of eHMI communication. They preferred seeking confirmation from co-located pedestrians in eHMI-related scenarios. P24 stated, *'We'll have a bit of reliance on other people because I don't really trust this automatic assistance in the car.'*

5.8.3 Vehicle Behaviour in a Multi-Pedestrian Scenario.

Participants frequently expressed feelings of confusion (n=15), particularly in scenarios where the AV stopped for another pedestrian but not for them. Such scenarios evoked feelings of unease and were often described using terms like *'concerning'* or *'weird'*. P8 articulated, *'From our perspective, the car should stop for both pedestrians, right? Like right here.'* The behaviour of the AV prompted questions about its detection mechanisms. P2 commented, *'I just feel maybe something went wrong because the vehicle didn't recognise me.'* Adding to this sentiment, P9 mentioned, *'I'm standing next to a tree, I'm uncertain if the car knows I'm there.'* P13 provided another perspective, noting that the vehicle seemed to operate on its own logic, choosing not to stop for them but opting to stop for another person shortly after. The reason behind such behaviour remained unclear to them.

Interestingly, these perceptions around the AV's stopping behaviour were seemingly diminished when considering the smartglasses concept. With a heightened focus on the signal, participants were less concerned by the AV stopping for someone else (n=5). P17 explained, *'I think it was less of an issue because I have my own personal device and all I'm waiting to see is a pedestrian crossing in front of me or not.'* This sense of personalisation was reflected in how participants were generally less concerned about the AV's actions in these AR-related scenarios (n=5).

5.8.4 Multi-Pedestrian VR Simulation.

Participants frequently appreciated the presence of a co-located pedestrian in the simulation (n=8). P17 provided insight into this sentiment by stating, *'I would've maybe found it a little bit sort of spooky and isolating, but to start the experience knowing that there's going to be another person and then just waving at the person. I actually felt like it was a little bit more friendly of an environment to be in some way.'*

The inclusion of a self-avatar was less impactful for participants when it was not directly visible to them (n=9). However, the most visible parts, namely the hands and legs, drew their attention (n=4). P20 mentioned, *'I noticed my shoes as I walked, making it feel more realistic.'* The embodiment in the simulation appeared to enhance the sense of immersion for some (n=5). As P14 stated, *'I have done it before where it's just kind of like my eyes in the world and it feels like I'm looking in, whereas this is the first one I've used where I really felt like I was in the world.'*

There were consistent reports from participants about issues related to avatar performance, with concerns mainly about its appearance and movement (n=14). P20 highlighted, *'my hands are pretty fake, because they looked like the doll's hand.'* P5 observed some irregularities, mentioning, *'There were some weird things going on with their legs. But that was only noticeable when I stopped.'* However, it is important to note that all participants mentioned that these avatar issues did not affect their decisions during crossings. P11 stated, *'I was generally able to interpret the avatar's movement direction. It was reasonably intuitive.'*

6 DISCUSSION

6.1 Clarity of Communication Among Interface Conditions (RQ1)

Overall, our results indicate that all eHMIs outperform the *Baseline* in crossing performance, specific intent communication, and interaction experiences. However, the *Vehicle* condition tends to cause misinterpretations. In contrast, the *Infrastructure* and *Pedestrian* conditions demonstrate similar yet better performance to the *Vehicle* across evaluated metrics, though not always significantly.

6.1.1 Crossing Performance.

Regarding crossing performance, the *Vehicle* condition has a significantly higher CIT compared to the *Infrastructure* (in 3 out of 4 scenarios) and the *Pedestrian* (in 1 out of 4 scenarios). This result could be partly attributed to the late activation strategy of the eHMI implemented in the *Vehicle* concept, with a

20-m activation as opposed to the 40-m activation in the *Infrastructure* and *Pedestrian* concepts. However, even with a compromised efficiency, the risk of misunderstanding AV signals still persists, as evidenced by qualitative data. This includes two instances where participants were misled by the eHMI's inward-sweeping animation. They began to cross prematurely and found the AV not stopping for them. Compared to the study findings by Dey et al. [19], these near-misses confirm the critical importance of focusing the yielding message of an eHMI to a recipient, especially in *one-other* scenarios. At the same time, they suggest that merely reducing willingness to cross in vehicle-based eHMI concepts may not be sufficient to ensure pedestrian safety. Notably, in the same high-risk scenario, the *Infrastructure* condition significantly lowered the CIT compared to the *Baseline* and *Vehicle* conditions, while also eliminating potential misunderstandings. This finding aligns with the conclusion by Holländer et al. [27] that the smart curbstones concept enhances the efficiency of traffic flow and is safer to use.

6.1.2 Specific Intent Communication. In the *Pedestrian* condition, participants experienced a strong sense of being directly communicated with, although this feeling was not significantly greater than in the *Vehicle* and *Infrastructure* conditions. This targeted communication capability is attributed to the personal nature of AR glasses. According to Dey et al. [17]'s taxonomy, this feature is highly scalable, meaning it can potentially address an unlimited number of road users. The *Vehicle* condition performed better than the *Baseline* in addressing individual pedestrians; however, there were three non-crossing decisions related to this condition. These instances demonstrate that targeted communication also has the potential to create ambiguity or exclusion among other road users who are not the primary targets of the communication. For example, in the case of P19, the decision not to cross when the AV communicated with another pedestrian illustrates how targeted communication can inadvertently signal to nearby pedestrians that they are not being acknowledged or considered. The interview data are in line with these findings, and concerns were noted in both scenarios where AV stopped for others (i.e., one-other, two-other).

6.1.3 Interaction Experiences. Regarding the Workload measure, eHMIs associated with pedestrian and infrastructure communications outperformed the *Baseline*. This suggests that such communication modes might be inherently more intuitive, offering pedestrians immediate clarity rather than requiring them to infer intentions. This immediacy is particularly beneficial in multi-pedestrian scenarios, where the presence of another individual often makes participants more hesitant to cross, indicating a heightened level of uncertainty compared to regular situations [71]. The question of why the *Vehicle* did not provide a significant cognitive advantage over the *Baseline* remains intriguing. Several factors appear to contribute to this outcome. First, the design of the light band employed a mass communication strategy. Consequently, pedestrians could only discern the intention of the AV to stop for them once it came to a complete stop and projected a crossing. This method kept participants in a state of uncertainty until the last possible moment, paralleling the uncertainty inherent in the *Baseline* condition. Second, eHMI light band concepts in general were found to significantly increase a pedestrian's willingness to cross [19, 71] compared to no eHMI. As a result, upon seeing the sweeping animation of the light band, participants expected that the AV would stop for them. This led to potential misunderstandings or frustration when it did not. In contrast, the study by Colley et al. [12] demonstrated a different outcome. Their design introduced the projection at an earlier point—when the AV was still 10 m away from the pedestrian it intended to yield for. This early signalling provided pedestrians with advance notice of where or for whom the AV intended to stop, significantly reducing cognitive workload compared to their baseline. Their findings highlight the impact of timely communication on reducing pedestrian uncertainty and aligning their expectations with the AV's actions.

The Trust measure presented varied findings. For the overall trust scores, only the *Pedestrian* performed better than the *Baseline*. When considering the Reliability/Competence subscale, again, only the *Pedestrian* exceeded the *Baseline* results. This is an interesting result, given the mixed findings from prior studies on AR applications within AV–pedestrian interactions. Tran et al. [67] found that their augmented AR crosswalk—a conformal AR design rooted in the environment—resulted in greater distrust and reduced trust compared to a baseline scenario. However, difference in AR design should be noted. The AR design employed in our study took inspiration from the Nudge HUD concept by Tabone et al. [61]. This design, which follows user eye movements, was found to provide a seamless experience that boosted user experience in making crossing decisions. Our qualitative analysis observed that information from the pedestrian-based eHMI remained visible even when their attention was diverted elsewhere. It suggests that the specific design and execution of AR interfaces can have varied impacts on trust. The trust findings also reveal a subtle discrepancy between perceived reliability and the source of communication. While the qualitative feedback indicated that communication directly from the AV was seen as reliable, the *Pedestrian* concept, which does not directly originate from the AV, received higher trust ratings. A potential interpretation for this could be that pedestrians value the clarity and immediacy of the information, perhaps even over its origin. Furthermore, qualitative feedback indicated that participants perceived AR glasses as being more seamlessly integrated with the broader smart city infrastructure and its users.

A noteworthy observation regarding Trust is that for the Understandability/Predictability subscale, no eHMI demonstrated a marked improvement over the *Baseline*. This suggests that while certain eHMIs might be perceived as more reliable, their ability to convey intentions in an understandable and predictable manner may need further attention. The vehicle's behaviour in multi-pedestrian scenarios could be a key contributing factor to this outcome. Qualitative insights highlighted participant confusion in situations where the AV chose to stop for one pedestrian over another. We suggest that future research should explore ways to enhance the predictability and comprehensibility of AV behaviour for pedestrians. Aside from AV behaviour, another aspect worth considering is the operation mode of the *Infrastructure* and *Pedestrian* concepts. The connections between these concepts and the vehicles were not entirely clear to the participants, prompting them to rely on previous experiences for interpretation. To address this issue, public education should be considered. Alternatively, the design could be made more transparent in its operations. For instance, the wearable AR concept developed by Tran et al. [67] informs the user with a message stating, 'Please wait. Communicating with oncoming vehicles,' enhancing user understanding of the technology's function.

In terms of UX, the *Pedestrian* condition has the highest mean ratings, though not significantly higher than the *Vehicle* and *Infrastructure* concepts. This trend is also reflected in the Preference ranking, where the *Pedestrian* condition is the most preferred, but statistical significance was not achieved when compared to other eHMI conditions. This might mean that, while each eHMI had its own unique characteristics, their overall impact on pedestrian overall experience was comparable. Future research should consider the strengths and weaknesses of these eHMIs as identified in this study. This will aid in refining them and in exploring particular scenarios where one eHMI may prove more effective than others.

6.2 Co-Located Pedestrian Influence (RQ2)

Prior research has highlighted the risks of ambiguity stemming from non-targeted AV communication [19, 20, 71]. Building on this, our findings underscore the role that co-located pedestrians play in individual pedestrian decisions. The qualitative results indicate a broad range of pedestrian behaviours, ranging from those who remain steadfast in their autonomy to those influenced, either subtly or overtly, by the presence of co-located pedestrians. In scenarios where AV communication lacks specificity, participants often turned to alternative information sources, particularly the actions of co-located pedestrians who were acknowledged by AVs, to feel safer when initiating their movements. This behaviour contrasts with dynamics observed in pedestrian groups, where understanding of eHMIs is typically influenced indirectly through a subconscious process of imitation rather than attentive observation [76]. In such group settings, individuals may unconsciously mimic the actions of others, a tendency driven by the psychological need to seek social proof in uncertain situations [9, 22]. Conversely, our study suggests that in co-located settings, participants actively use the behaviour cues of nearby pedestrians not merely to conform to social norms but to gather additional information for making well-judged decisions, suggesting a more rational decision-making process.

Qualitative findings also found that the clearer the instructions provided by the AVs, the lesser the influence of co-located pedestrians on individual decisions. The shift from no particular guidance in the *Baseline*, advances to basic crossing suggestions in the *Vehicle*, and ends with safe/unsafe crossing indicators in the *Infrastructure* and *Pedestrian* concepts seems to influence pedestrians to trust the technology more and rely on their co-located pedestrians less. This observation aligns with research on pedestrian group dynamics and eHMIs, which suggests that the influence of other pedestrians on crossing decisions diminishes as individuals gain familiarity with eHMIs [76]. However, this shift towards trust in technology, while beneficial in enhancing predictability in AV–pedestrian interactions, comes with its own set of concerns.

(1) Technology Reliance: In line with our findings, data from Tabone et al. [61] indicate that participants in the Nudge HUD concept often did not glance at the AV (in 70 out of 360 trials). This suggests that as communication from AVs becomes clearer, pedestrians might direct their attention more towards the instructions and less towards the vehicle, co-located pedestrians, and other environmental cues.

(2) Pedestrian Autonomy in Decision-Making: A segment of participant feedback touched on the concept of pedestrian autonomy. As instructions from the AVs became more defined, some pedestrians felt a change in their decision-making process. It is worth noting that individual reactions to the presence of crossing instructions varied, with some appreciating the clarity, while others expressed a desire for more personal judgment in the scenarios. This loss of autonomy can lead to reduced confidence in real-world scenarios where the technology might be absent or malfunctioning.

6.3 Limitations and Future Work

One limitation of our study is the small sample size; with only 25 participants, the findings may lack the statistical power necessary to detect smaller effects. Additionally, our study did not include scenarios in which a pedestrian crosses the road alone, devoid of any additional pedestrian in the scene, such as the Solo conditions studied by Colley et al. [12]. Including such conditions could have provided unique insights into individual decision-making processes and crossing behaviours in the absence of co-located pedestrian influence.

The research was conducted in VR, primarily for safety reasons and to reduce the development costs associated with creating futuristic prototypes such as smart infrastructure and wearable AR [50, 59, 64, 65]. While this approach offers significant benefits, it may not fully capture the complexities of real-world settings. The traffic was simplified, featuring only one approaching vehicle, and lacked background social activities, potentially reducing the environmental realism and participants' sense of presence. Moreover, the virtual setting did not account for environmental variables that could significantly impact the visibility of design concepts. For example, the vehicle-based projection may be compromised bright sunlight [12, 19, 51] and obstructed road surfaces. Future research should consider more complex traffic simulations and utilise real-world representations (e.g., 360-degree capture of real world) to uncover interface issues under more natural conditions [26].

Through the social presence score, our simulator demonstrated that participants perceived co-located pedestrians as genuine social actors. However, the avatars' movements, based on inverse kinematics and Meta Quest 2 tracking data, did not perfectly match real human movements. The study also lacked gaze and eye-tracking, which could have provided deeper insights into how pedestrians engage visually with AVs, eHMIs, and other pedestrians. Although the current experimental setup with two pedestrians does not constitute a scenario with truly multiple pedestrians, it serves as a methodical step to gradually study the social effects in AV—pedestrian interactions. With practical lessons learned from this study, we plan to enhance the simulator to include a larger number of pedestrians joining remotely from different locations, support improved representations of their virtual embodiments, and collect more objective behavioural data.

Finally, in selecting representatives for each communication locus, we prioritised those with notable popularity and potential for scalable AV communication. Efforts were made to minimise differences by using common visual elements, such as colour schemes and zebra crossings. However, it is important to note that our findings may be somewhat specific to the chosen embodiments of these concepts, and may not fully generalise to the broader categories of communication loci they represent.

7 CONCLUSION

In response to the call for more research on eHMI scalability, our exploration of AV–pedestrian interaction in multi-pedestrian scenarios highlighted the interplay between technology, individual decisions, and the influence of co-located pedestrians. The introduction of eHMIs offers promising avenues for clearer communication, yet no single concept consistently surpassed others in multi-pedestrian scenarios, suggesting that there may not be a one-size-fits-all solution that excels universally. With the strengths and weaknesses of each communication locus and its corresponding concept outlined, we provide foundational insights for future development and considerations for their appropriate application.

The presence of co-located pedestrians introduces uncertainty in AV behaviour, often prompting pedestrians to actively observe their peers as input for decision-making in ambiguous situations. However, this behaviour diminishes with clearer and more targeted AV communications. As we look towards a future with an increased AV presence in urban landscapes, it is crucial to strike a balance between technological guidance and individual autonomy. The ultimate goal remains to ensure safety while preserving human-centric interactions and supporting individual decision-making.

ACKNOWLEDGMENTS

This research is supported by an Australian Government Research Training Program (RTP) Scholarship and through the Australian Research Council (ARC) Discovery Project DP220102019, Shared-space interactions between people and autonomous vehicles. We thank all the participants for taking part in this research. We extend our gratitude to the anonymous reviewers for their insightful comments and suggestions.

To support open science initiatives, we have initially made the scripts from the multi-pedestrian VR simulator available on GitHub, with plans to release the entire project in the near future: https://github.com/minhtramtran/multiped-sim.

REFERENCES

[1] A. Aristidou, J. Lasenby, Y. Chrysanthou, and A. Shamir. 2018. Inverse Kinematics Techniques in Computer Graphics: A Survey. *Computer Graphics Forum* 37, 6 (2018), 35–58. https://doi.org/10.1111/cgf.13310

[2] Ronald T Azuma. 2019. The road to ubiquitous consumer augmented reality systems. *Human Behavior and Emerging Technologies* 1, 1 (2019), 26–32.

[3] Pavlo Bazilinskyy, Dimitra Dodou, and Joost de Winter. 2020. External Human-Machine Interfaces: Which of 729 Colors Is Best for Signaling 'Please (Do not) Cross'?. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 3721–3728. https://doi.org/10.1109/SMC42975.2020.9282998

[4] Pavlo Bazilinskyy, Lars Kooijman, Dimitra Dodou, and Joost CF de Winter. 2020. Coupled simulator for research on the interaction between pedestrians and (automated) vehicles. In *Driving Simulation Conference Europe. Antibes, France.*

[5] Jelte E Bos, Sjoerd C de Vries, Martijn L van Emmerik, and Eric L Groen. 2010. The effect of internal and external fields of view on visually induced motion sickness. *Applied ergonomics* 41, 4 (2010), 516–521.

[6] James V Bradley. 1958. Complete counterbalancing of immediate sequential effects in a Latin square design. *J. Amer. Statist. Assoc.* 53, 282 (1958), 525–528.

[7] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[8] Henry Chen, Robin Cohen, Kerstin Dautenhahn, Edith Law, and Krzysztof Czarnecki. 2020. Autonomous Vehicle Visual Signals for Pedestrians: Experiments and Design Recommendations. In *2020 IEEE Intelligent Vehicles Symposium (IV)*. Ieee, New Jersey, NJ, USA,

1819–1826. https://doi.org/10.1109/iv47402.2020.9304628

[9] Robert B Cialdini. 2009. *Influence: Science and practice*. Vol. 4. Pearson education Boston.

[10] Mark Colley, Elvedin Bajrovic, and Enrico Rukzio. 2022. Effects of pedestrian behavior, time pressure, and repeated exposure on crossing decisions in front of automated vehicles equipped with external communication. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–11.

[11] Mark Colley, Jan Henry Belz, and Enrico Rukzio. 2021. Investigating the Effects of Feedback Communication of Autonomous Vehicles. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) *(AutomotiveUI '21)*. Association for Computing Machinery, New York, NY, USA, 263–273. https://doi.org/10.1145/3409118.3475133

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

[13] Mark Colley and Enrico Rukzio. 2020. A Design Space for External Communication of Autonomous Vehicles. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) *(AutomotiveUI '20)*. Association for Computing Machinery, New York, NY, USA, 212–222. https://doi.org/10.1145/3409120.3410646

[14] Mark Colley, Marcel Walch, and Enrico Rukzio. 2020. Unveiling the Lack of Scalability in Research on External Communication of Autonomous Vehicles. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3382865

[15] Microsoft Corporation. 2023. Microsoft Mesh for Microsoft Teams. https://www.microsoft.com/en-us/microsoft-teams/microsoft-mesh

[16] Joost de Winter and Dimitra Dodou. 2022. External Human-Machine Interfaces: Gimmick or Necessity. *Delft University of Technology, Delft, The Netherlands* (2022).

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

[18] Debargha Dey, Azra Habibovic, Bastian Pfleging, Marieke Martens, and Jacques Terken. 2020. Color and Animation Preferences for a Light Band EHMI in Interactions Between Automated Vehicles and Pedestrians. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376325

[19] Debargha Dey, Arjen van Vastenhoven, Raymond H Cuijpers, Marieke Martens, and Bastian Pfleging. 2021. Towards scalable eHMIs: Designing for AV-VRU communication beyond one pedestrian. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 274–286.

[20] André Dietrich, J-H Willrodt, Karolin Wagner, and Klaus Bengler. 2019. Projection-based external human machine interfaces-enabling interaction between automated vehicles and pedestrians. In *Proceedings of the DSC 2018 Europe VR*.

[21] Stefanie M. Faas and Martin Baumann. 2019. Light-Based External Human Machine Interface: Color Evaluation for Self-Driving Vehicle and Pedestrian Interaction. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 63, 1 (2019), 1232–1236. https://doi.org/10.1177/1071181319631049

[22] Milad Haghani and Majid Sarvi. 2016. Pedestrian crowd tactical-level decision making during emergency evacuations. *Journal of Advanced Transportation* 50, 8 (2016), 1870–1895.

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

[24] Marc Hesenius, Ingo Börsting, Ole Meyer, and Volker Gruhn. 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*. 261–268.

[25] Marius Hoggenmüller and Martin Tomitsch. 2019. Enhancing Pedestrian Safety through In-Situ Projections: A Hyperreal Design Approach. In *Proceedings of the 8th ACM International Symposium on Pervasive Displays* (Palermo, Italy) *(PerDis '19)*. Association for Computing Machinery, New York, NY, USA, Article 29, 2 pages. https://doi.org/10.1145/3321335.3329682

[26] Marius Hoggenmüller, Martin Tomitsch, Luke Hespanhol, Tram Thi Minh Tran, Stewart Worrall, and Eduardo Nebot. 2021. Context-based interface prototyping: Understanding the effect of prototype representation on user feedback. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[27] Kai Holländer, Marius Hoggenmüller, Romy Gruber, Sarah Theres Völkel, and Andreas Butz. 2022. Take It to the Curb: Scalable Communication Between Autonomous Cars and Vulnerable Road Users Through Curbstone Displays. *Frontiers in Computer Science* 4 (2022), 24.

[28] Kai Holländer, Andy Krüger, and Andreas Butz. 2020. Save the Smombies: App-Assisted Street Crossing. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–11.

[29] Ahmed Hussein, Fernando Garcia, Jose Maria Armingol, and Cristina Olaverri-Monreal. 2016. P2V and V2P communication for pedestrian warning on the basis of autonomous vehicles. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*.

IEEE, 2034–2039.

[30] Yuanyuan Jiang, Elizabeth E. O'Neal, Pooya Rahimian, Junghum Paul Yon, Jodie M. Plumert, and Joseph K. Kearney. 2019. Joint Action in a Virtual Environment: Crossing Roads with Risky vs. Safe Human and Agent Partners. *IEEE Transactions on Visualization and Computer Graphics* 25, 10 (2019), 2886–2895. https://doi.org/10.1109/TVCG.2018.2865945

[31] Dongsik Jo, Kangsoo Kim, Gregory F Welch, Woojin Jeon, Yongwan Kim, Ki-Hong Kim, and Gerard Jounghyun Kim. 2017. The impact of avatar-owner visual similarity on body ownership in immersive virtual reality. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*. 1–2.

[32] Philip Joisten, Ziyu Liu, Nina Theobald, Andreas Webler, and Bettina Abendroth. 2021. Communication of Automated Vehicles and Pedestrian Groups: An Intercultural Study on Pedestrians' Street Crossing Decisions. In *Proceedings of Mensch Und Computer 2021*. Association for Computing Machinery, New York, NY, USA, 49–53. https://doi.org/10.1145/3473856.3474004

[33] Amir Hossein Kalantari, Yue Yang, Jorge Garcia de Pedro, Yee Mun Lee, Anthony Horrobin, Albert Solernou, Christopher Holmes, Natasha Merat, and Gustav Markkula. 2023. Who goes first? a distributed simulator study of vehicle–pedestrian interaction. *Accident Analysis & Prevention* 186 (2023), 107050.

[34] Joseph K Kearney, Jodie M Plumert, Chris Schwarz, Stephen Baek, Elizabeth O'Neal, Wanxin Wang, Daniel V McGehee, Deborah Beckwin, et al. 2020. *Learning About Driver and Pedestrian Behaviors Through Connected Simulation Technology*. Technical Report. United States. Federal Highway Administration.

[35] Hayeon Kim, Jinhyung Park, and In-Kwon Lee. 2023. "To be or Not to be Me?": Exploration of Self-Similar Effects of Avatars on Social Virtual Reality Experiences. *IEEE Transactions on Visualization and Computer Graphics* 29, 11 (2023), 4794–4804. https://doi.org/10.1109/TVCG.2023.3320240

[36] Max Kinateder and William H Warren. 2016. Social influence on evacuation behavior in real and virtual environments. *Frontiers in Robotics and AI* 3 (2016), 43.

[37] Lars Kooijman, Riender Happee, and Joost C. F. de Winter. 2019. How Do eHMIs Affect Pedestrians' Crossing Behavior? A Study Using a Head-Mounted Display Combined with a Motion Suit. *Information* 10, 12 (2019). https://doi.org/10.3390/info10120386

[38] Moritz Körber. 2019. Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*. Springer International Publishing, Cham, 13–30.

[39] Sabrina Lakhdhir, Sowmya Somanath, and Ehud Sharlin. 2023. Wearing Awareness: Designing Pedestrian-Wearables for Interactions with Autonomous Vehicles. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 316, 8 pages. https://doi.org/10.1145/3544549.3585655

[40] Mirjam Lanzer, Ina Koniakowsky, Mark Colley, and Martin Baumann. 2023. Interaction Effects of Pedestrian Behavior, Smartphone Distraction and External Communication of Automated Vehicles on Crossing and Gaze Behavior. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.

[41] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings 4*. Springer, 63–76.

[42] Andreas Löcken, Carmen Golling, and Andreas Riener. 2019. How should automated vehicles interact with pedestrians? A comparative analysis of interaction concepts in virtual reality. In *Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications*. 262–274.

[43] Alexandra Loew, Julia Graefe, Lukas Heil, Anne Guthardt, Annika Boos, André Dietrich, and Klaus Bengler. 2022. Go Ahead, Please!—Evaluation of External Human—Machine Interfaces in a Real-World Crossing Scenario. *Frontiers in Computer Science* 4 (2022). https://doi.org/10.3389/fcomp.2022.863072

[44] Karthik Mahadevan, Elaheh Sanoubari, Sowmya Somanath, James E. Young, and Ehud Sharlin. 2019. AV-Pedestrian interaction design using a pedestrian mixed traffic simulator. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(Dis '19)*. Association for Computing Machinery, New York, NY, USA, 475–486. https://doi.org/10.1145/3322276.3322328

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

[46] Guido Makransky, Lau Lilleholt, and Anders Aaby. 2017. Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior* 72 (2017), 276–285.

[47] Jeehan Malik, Morgan N. Di Napoli Parr, Jessica Flathau, Hanxi Tang, Joseph K. Kearney, Jodie M. Plumert, and Kyle Rector. 2021. Determining the effect of smartphone alerts and warnings on street-crossing behavior in non-mobility-impaired older and younger adults. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.

[48] Mercedes-Benz. 2023. Mercedes-Benz Receives Approvals for Turquoise-Coloured Automated Driving Marker Lights in California and Nevada. Online. https://media.mercedes-benz.be/fr/mercedes-benz-receives-approvals-for-turquoise-coloured-automated-driving-marker-lights-in-california-and-nevada/ Accessed: 2024-04-16.

[49] Meta. 2023. Meta Horizon Worlds on Meta Quest. https://www.meta.com/en-gb/experiences/2532035600194083/

- [50] Alexandre M Nascimento et al. 2019. The role of virtual reality in autonomous vehicles' safety. In *2019 IEEE Int. Conf. Artif. Intell. Virtual Reality*. 50–507.
- [51] Trung Thanh Nguyen, Kai Holländer, Marius Hoggenmueller, Callum Parker, and Martin Tomitsch. 2019. Designing for Projection-Based Communication between Autonomous Vehicles and Pedestrians. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 284–294. https://doi.org/10.1145/3342197.3344543
- [52] Nissan. 2015. Nissan IDS Concept: Nissan's vision for the future of EVs and autonomous driving. https://europe.nissannews.com/en-GB/releases/release-139047
- [53] F Gabriele Pratticò, Fabrizio Lamberti, Alberto Cannavò, Lia Morra, and Paolo Montuschi. 2021. Comparing State-of-the-Art and Emerging Augmented Reality Interfaces for Autonomous Vehicle-to-Pedestrian Communication. *IEEE Transactions on Vehicular Technology* 70, 2 (2021), 1157–1168.
- [54] Amir Rasouli and John K Tsotsos. 2019. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. *IEEE transactions on intelligent transportation systems* 21, 3 (2019), 900–918.
- [55] Andreas Riegler, Andreas Riener, and Clemens Holzmann. 2020. A Research Agenda for Mixed Reality in Automated Vehicles. In *Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia* (Essen, Germany) *(MUM '20)*. Association for Computing Machinery, New York, NY, USA, 119–131. https://doi.org/10.1145/3428361.3428390
- [56] Alexandros Rouchitsas and Håkan Alm. 2019. External human–machine interfaces for autonomous vehicle-to-pedestrian communication: A review of empirical work. *Frontiers in psychology* 10 (2019), 2757.
- [57] Ehsan Sadraei et al. 2020. Vehicle-pedestrian interaction: A distributed simulation study. In *Proc. 19th Driving Simul. Conf.* Antibes, France, 145–154.
- [58] SAE. 2019. *Automated Driving System (ADS) Marker Lamp.* SAE Recommended Practice J3134_201905. Society of Automotive Engineers, Warrendale, PA. 10 pages. https://doi.org/10.4271/J3134_201905
- [59] Sonja Schneider and Klaus Bengler. 2020. Virtually the same? Analysing pedestrian behaviour by means of virtual reality. *Transp. Res. Part F: Traffic Psychol. Behav.* 68 (2020), 231–256.
- [60] Robin Connor Schramm, Markus Sasalovici, Axel Hildebrand, and Ulrich Schwanecke. 2023. Assessing Augmented Reality Selection Techniques for Passengers in Moving Vehicles: A Real-World User Study. In *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ingolstadt, Germany) *(AutomotiveUI '23)*. Association for Computing Machinery, New York, NY, USA, 22–31. https://doi.org/10.1145/3580585.3607152
- [61] Wilbert Tabone, Riender Happee, Jorge García, Yee Mun Lee, Maria Luce Lupetti, Natasha Merat, and Joost de Winter. 2023. Augmented reality interfaces for pedestrian-vehicle interactions: An online study. *Transportation research part F: traffic psychology and behaviour* 94 (2023), 170–189.
- [62] Wilbert Tabone, Yee Mun Lee, Natasha Merat, Riender Happee, and Joost de Winter. 2021. Towards Future Pedestrian-Vehicle Interactions: Introducing Theoretically-Supported AR Prototypes. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) *(AutomotiveUI '21)*. Association for Computing Machinery, New York, NY, USA, 209–218. https://doi.org/10.1145/3409118.3475149
- [63] Kai Tian, Gustav Markkula, Chongfeng Wei, Yee Mun Lee, Ruth Madigan, Toshiya Hirose, Natasha Merat, and Richard Romano. 2023. Deconstructing Pedestrian Crossing Decisions in Interactions With Continuous Traffic: An Anthropomorphic Model. *IEEE Transactions on Intelligent Transportation Systems* (2023).
- [64] Tram Thi Minh Tran, Callum Parker, Marius Hoggenmüller, Luke Hespanhol, and Martin Tomitsch. 2023. Simulating wearable urban augmented reality experiences in vr: Lessons learnt from designing two future urban interfaces. *Multimodal Technologies and Interaction* 7, 2 (2023), 21.
- [65] Tram Thi Minh Tran, Callum Parker, and Martin Tomitsch. 2021. A Review of Virtual Reality Studies on Autonomous Vehicle–Pedestrian Interaction. *IEEE Transactions on Human-Machine Systems* 51, 6 (2021), 641–652. https://doi.org/10.1109/THMS.2021.3107517
- [66] Tram Thi Minh Tran, Callum Parker, and Martin Tomitsch. 2023. Scoping Out the Scalability Issues of Autonomous Vehicle-Pedestrian Interaction. In *15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ingolstadt, Germany) *(AutomotiveUI '23)*. Association for Computing Machinery, New York, NY, USA, 167–177. https://doi.org/10.1145/3580585.3607167
- [67] Tram Thi Minh Tran, Callum Parker, Yiyuan Wang, and Martin Tomitsch. 2022. Designing wearable augmented reality concepts to support scalability in autonomous vehicle–pedestrian interaction. *Frontiers in Computer Science* (2022), 39.
- [68] Umbrellium. 2017. Case study - make roads safer, more responsive & dynamic. https://umbrellium.co.uk/case-studies/south-london-starling-cv/
- [69] Rutger Verstegen, Debargha Dey, and Bastian Pfleging. 2021. CommDisk: A Holistic 360° EHMI Concept to Facilitate Scalable, Unambiguous Interactions between Automated Vehicles and Other Road Users. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) *(AutomotiveUI '21 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 132–136. https://doi.org/10.1145/3473682.3480280

[70] Thomas Waltemate, Dominik Gall, Daniel Roth, Mario Botsch, and Marc Erich Latoschik. 2018. The Impact of Avatar Personalization and Immersion on Virtual Body Ownership, Presence, and Emotional Response. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (2018), 1643–1652. https://doi.org/10.1109/TVCG.2018.2794629

[71] Marc Wilbrink, Manja Nuttelmann, and Michael Oehl. 2021. Scaling up automated vehicles' eHMI communication designs to interactions with multiple pedestrians–putting eHMIs to the test. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 119–122.

[72] Working Party on Automated/Autonomous and Connected Vehicles (GRVA). 2022. *Minutes of the 18th GRVA Session*. Meeting Report WP.29-188-10. United Nations Economic Commission for Europe. https://unece.org/sites/default/files/2022-11/WP.29-188-10e.pdf Accessed: 2024-04-16.

[73] Xinzhou Wu, Radovan Miucic, Sichao Yang, Samir Al-Stouhi, James Misener, Sue Bai, and Wai-hoi Chan. 2014. Cars talk to phones: A DSRC based vehicle-pedestrian safety system. In *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*. IEEE, 1–7.

[74] Putra AR Yamin and Jaehyun Park. 2023. Where Did That Come From? The Effects of eHMI Approach Angle of an Automated Personal Mobility Vehicle (APMV) on Pedestrian Crossing Behavior. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 203–208.

[75] Yue Yang, Amir Hossein Kalantari, Yee Mun Lee, Albert Solernou, Gustav Markkula, and Natasha Merat. 2023. A Distributed Simulation Study to Examine Vehicle – Pedestrian Interactions. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ingolstadt, Germany) *(AutomotiveUI '23 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 327–329. https://doi.org/10.1145/3581961.3609852

[76] Xiaoyuan Zhao, Xiaomeng Li, Andry Rakotonirainy, Samira Bourgeois-Bougrine, Zishuo Zhu, and Patricia Delhomme. 2024. Crossing roads in a social context: How behaviors of others shape pedestrian interaction with automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour* 102 (2024), 88–106. https://doi.org/10.1016/j.trf.2024.02.008