Designing Wearable Augmented Reality Concepts to Support Scalability in Autonomous Vehicle-Pedestrian Interaction

Tram Thi Minh Tran*, Callum Parker, Yiyuan Wang and Martin Tomitsch

Design Lab, Sydney School of Architecture, Design and Planning, The University of Sydney, Sydney, NSW, Australia

Wearable augmented reality (AR) offers new ways for supporting the interaction between autonomous vehicles (AVs) and pedestrians due to its ability to integrate timely and contextually relevant data into the user's field of view. This article presents novel wearable AR concepts that assist crossing pedestrians in multi-vehicle scenarios where several AVs frequent the road from both directions. Three concepts with different communication approaches for signaling responses from multiple AVs to a crossing request, as well as a conventional pedestrian push button, were simulated and tested within a virtual reality environment. The results showed that wearable AR is a promising way to reduce crossing pedestrians' cognitive load when the design offers both individual AV responses and a clear signal to cross. The willingness of pedestrians to adopt a wearable AR solution, however, is subject to different factors, including costs, data privacy, technical defects, liability risks, maintenance duties, and form factors. We further found that all participants favored sending a crossing request to AVs rather than waiting for the vehicles to detect their intentions—pointing to an important gap and opportunity in the current AV-pedestrian interaction literature.

Keywords: autonomous vehicles, vehicle-to-pedestrian communication, external human–machine interfaces, user-initiated communication, vulnerable road users, wearable augmented reality, scalability

1. INTRODUCTION

The ability of autonomous vehicles (AVs) to effectively interact with vulnerable road users (VRUs), such as pedestrians, is crucial to ensuring safe operations and public confidence. While pedestrians mainly rely on implicit cues (e.g., motion and motor sounds) from a vehicle to interpret its intention (Risto et al., 2007; Moore et al., 2019), explicit signals from a driver, including verbal exchanges, eye contact, and hand gestures, help resolve impasses and instill trust in interactions (Rasouli and Tsotsos, 2019). Once humans relinquish control to an AV, these informal signals may become less prevalent or possibly disappear altogether. External human-machine interfaces (eHMIs) (Dey et al., 2020a) are currently being investigated as a possible way to compensate for the lack of driver cues, allowing for intention transparency, which is a desirable quality in almost every intelligent system (Zileli et al., 2019).

In order to understand key factors influencing pedestrian behavior and experiences, most external communication research has evaluated eHMIs in the fundamental traffic setting involving one pedestrian and one vehicle (Colley et al., 2020b). However, for eHMIs to become an effective

mediator in real-world traffic situations, it is critical for external communication research to take into account scalability factors (i.e., vehicular and pedestrian traffic volumes) and their associated challenges. For example, pedestrians may experience an increased cognitive load when interpreting signals from multiple AVs (Mahadevan et al., 2018; Dey et al., 2020a) or mistakenly believe a message intended for another is directed to them (Dey et al., 2021).

One promising solution to the scalability issues is incorporating augmented reality (AR). This technology has been explored in the automobile industry to improve driving safety and comfort (Riegler et al., 2021). In-car AR, such as heads-up and windshield displays, offer diverse opportunities to aid navigation, highlight potential hazards, and allow for a shared perception between a driver and an automated driving system (Wiegand et al., 2019). The application of AR outside of vehicles to assist AV-pedestrian interaction is also of increasing interest in academia (Tabone et al., 2021a). As with smartphones, the personal nature of wearable AR[1] allows their connected eHMI concepts to address an unlimited number of road users simultaneously with notable precision and resolution (Dey et al., 2020a). In addition, tailored communication based on user preferences and characteristics may contribute to eHMIs becoming more inclusive. Notably, AR has been investigated as an accessibility tool for visually impaired people (Coughlan and Miele, 2017). Most significantly, wearable AR enables digital content to be displayed within the physical environment, allowing users to retain situational awareness and react rapidly to safety alerts (Tong et al., 2021).

Various AR concepts have been designed to convey road-crossing information (Hesenius et al., 2018; Prattico et al., 2021; Tabone et al., 2021b) and provide collision warnings to pedestrians (Tong et al., 2021). However, to our knowledge, no studies have been undertaken to date to evaluate AR concepts in a complex traffic setting where pedestrians must consider the intentions of several AVs in making crossing decisions. Our driving assumption is that a multi-vehicle situation necessitates the understanding of an appropriate communication approach to provide pedestrians with pertinent cues without overwhelming them. Furthermore, the literature has focused on determining the efficacy of various AR concepts in conveying AV intent rather than pedestrians' preferences for using wearable AR in daily interactions with AVs. Given the novel AR experiences and the shift away from using public crossing facilities and toward using personal devices, it is important to gauge pedestrians' acceptance of wearable AR solutions.

To address these research gaps, we designed three AR eHMI concepts with different ways to signal responses from multiple AVs: a visual cue on each vehicle, a visual cue that represents all vehicles, and both the aforementioned types of visual cues. We used virtual reality (VR) to simulate and test wearable AR prototypes against a pedestrian push button baseline. Our overall research goal was to answer the following research questions: **(RQ1)** *To what extent, if any, do pedestrians prefer using wearable AR to interact with AVs?* **(RQ2)** *How do different communication approaches influence pedestrians' perceived cognitive load and trust?*

Our study makes the following contributions: we (1) present novel AR eHMI concepts that assist the crossing of pedestrians in heavy traffic scenarios, (2) identify factors influencing pedestrians' preferences for wearable AR solutions, and (3) determine the effect of three distinct communication approaches on pedestrians' crossing experiences.

2. RELATED WORK

2.1. External Communication of AVs

The vast majority of car crashes are caused by human error (Treat et al., 1979; Hendricks et al., 2001); therefore, advanced driver assistance systems have been developed to assist drivers in a variety of driving tasks (e.g., active cruise control, collision warnings) or relieve them fully from driving. Without active drivers, future vehicles may be outfitted with additional interfaces that communicate clearly with pedestrians and other VRUs regarding their intentions and operating states. For instance, Waymo has submitted a patent stating that cars may inform pedestrians using "a physical signaling device, an electronic sign or lights, [or] a speaker for providing audible notifications" (Urmson et al., 2015). Meanwhile, Uber has further proposed using a virtual driver and on-road projections (Sweeney et al., 2018). Potential implementations of eHMI also include approaches in which communication messages are detached from the vehicles. The urban technology firm Umbrellium has prototyped an LED-based road surface capable of dynamically adapting its road markings to different traffic conditions to prioritize pedestrians' safety (Umbrellium, 2017). In addition, Telstra has trialed a technology enabling vehicles to deliver early-warning collision alerts to pedestrians *via* a smartphone (Cohda Wireless, 2017).

The locus of communication—*Vehicle*, *Infrastructure*, and *Personal Device*—is one of the key dimensions in the eHMI design space (Colley and Rukzio, 2020a). According to a review of 70 different design concepts from industry and academia, vehicle-mounted devices have accounted for the majority of research on the external communication of AVs thus far (Dey et al., 2020a). However, urban infrastructure and personal devices are promising alternatives for facilitating complex interactions involving multiple road users and vehicles due to their high scalability and communication resolution (Dey et al., 2020a).

2.2. Scalability

In the context of AV-pedestrian interaction, scalability refers to the ability of eHMIs to be employed in situations with a large number of vehicles and pedestrians without compromising on efficacy (Colley et al., 2020b). In this case, the communication relationship goes beyond the simple one-to-one encounters and includes one-to-many, many-to-one, and many-to-many interactions (Colley and Rukzio, 2020b). Although scalability research is still in its early stages (Colley et al., 2020b), potential

---

[1]In this article, we use the term *wearable AR* to refer to all types of near-eye displays regardless of their form factor. These displays include head-mounted AR devices (e.g., Microsoft HoloLens), monocular and binocular AR glasses (e.g., Google Glass), and contact lenses (e.g., Mojo Lens).

scaling limitations of eHMIs, including low communication resolution and information overload, have been noted in several research articles (Robert Jr, 2019; Dey et al., 2020a, 2021).

In terms of communication resolution, i.e., "the clarity of whom the message of an eHMI is intended" (Dey et al., 2020a), a message broadcasted to all road users in a vehicle's vicinity, e.g., from an on-vehicle LED display, might result in misinterpretation. This issue is particularly apparent when co-located road users have conflicting rights of way (Dey et al., 2020a), which may lead to confusion or even unfortunate outcomes in real-world traffic situations. Dey et al. (2021) tested four eHMI designs with two pedestrians and observed that non-specific yielding messages increased the participants' willingness to cross even when the vehicle was stopping for another person. To address this possible communication failure, Verstegen et al. (2021) prototyped a 360-degree disk-shaped eHMI featuring eyes and dots that acknowledge the presence of multiple (groups of) pedestrians. Other proposed alternatives include nomadic devices, the personal nature of which inherently enables targeted communication, and smart infrastructures (e.g., responsive road surfaces) (Dey et al., 2020a). However, more research is required to determine user acceptance and the (cost-) effectiveness of such solutions.

Information overload may occur when pedestrians are presented with an excessive number of cues. In the study by Mahadevan et al. (2018), a mixed interface of three explicit cues situated on the automobile, street infrastructure, and a pedestrian's smartphone was viewed as time-consuming and perplexing by many participants. Hesenius et al. (2018) reported a similar finding, where participants disliked the prototype that visualizes safe zones, navigation paths, and vehicle intents simultaneously. In the case of multiple AVs, an increase in the number of external displays was expected to impose a high cognitive load onto pedestrians (Robert Jr, 2019) and turn street crossing into "an analytical process" (Moore et al., 2019). According to Colley et al. (2020a), when multiple AVs communicate using auditory messages, pedestrians' perceived safety and cognitive load improve; however, it is uncertain whether the same observation can be made with visual messages. Our study aims to close this knowledge gap by examining three different approaches to displaying visual responses from multiple AVs.

2.3. Wearable AR Concepts

Globally, smartphone uptake has increased at a very swift pace. Together with advances in short-range communication technologies, the devices have been investigated for their potential to improve pedestrian safety, such as aiding individuals in crossing streets (Holländer et al., 2020; Malik et al., 2021) and providing collision alerts (Wu et al., 2014; Hussein et al., 2016). Smartphones' close proximity to users allows them to access reliable positioning data for collision estimations and deliver adaptive communication messages based on users' current phone activity (e.g., listening to music) (Liu et al., 2015).

Wearable AR, as the next wave of computing innovation, has similar advantages to smartphones. However, its ability to combine the virtual and real worlds enables a more compelling

and natural display of information and improved retention of situational awareness (Azuma, 2019). Context-aware and pervasive AR applications (Grubert et al., 2016) also present an opportunity to aid users in more diverse ways. They are envisioned to become smart assistants that can semantically understand the surrounding environment, monitor the user's current states (e.g., gaze and visual attention), and adjust to their situational needs (Starner et al., 1997; Azuma, 2019). This has led to a growing discussion on the application of wearable AR for AV-pedestrian communication. In a position paper where 16 scientific experts were interviewed, it was partially agreed that wearable AR might resolve scalability issues of AV-VRU interaction (Tabone et al., 2021a). Recent work has explored several different AR eHMI concepts but has yet to examine the scalability aspect. Tong and Jia (2019) designed an AR interface to warn pedestrians of oncoming vehicles while other studies have presented navigational concepts (Hesenius et al., 2018; Prattico et al., 2021) and theoretically-supported prototypes (Tabone et al., 2021b) offering crossing advice. Our study attempts to extend this body of work through an empirically based investigation of wearable AR design concepts in a multi-vehicle situation.

Currently, various technical issues exist that make it challenging to prototype and evaluate wearable AR interfaces outdoors (Billinghurst, 2021): (1) a narrow field of view (FOV) that covers only a portion of the human field of vision, limiting what users can see to a small window; (2) an unstable tracking system that is affected by a wide range of environmental factors (e.g., lighting, temperature, and movement in space); and (3) low visibility of the holograms in direct sunlight. For these reasons, we utilized VR simulations to overcome the shortcomings of wearable AR and the limitations of AV testing in the real world, following a similar approach to Prattico et al. (2021).

3. DESIGN PROCESS

3.1. Crossing Scenario

Similar to most studies on AV-pedestrian interaction, we selected an ambiguous traffic situation, i.e., a midblock location without marked crosswalks or traffic signals, requiring pedestrians to cross with caution and be vigilant of oncoming vehicles. To assess the design concept's scalability, the crossing scenario featured many vehicles driving in both directions on a two-way street. This situation is prevalent in urban traffic, typically requiring pedestrians to estimate the time-to-arrival of vehicles and select a safe gap to cross. However, the ability to correctly assess the speed and distance of approaching cars varies with different environmental conditions and across demographic groups (Rasouli and Tsotsos, 2019). Inaccurate judgments may lead to unsafe crossing decisions, causing pedestrian conflicts with vehicular traffic. On the premise that not all road users can chart their best course of action, we sought to create a design concept to aid their crossing decisions.

3.2. Design Concepts

Our wearable AR concepts were inspired by the widely used pedestrian push button, which enables pedestrians to

request a crossing phase. The buttons are typically installed at locations with intermittent pedestrian volumes, where an automatic pedestrian walk phase has not been implemented. The installment is intended to improve vehicle mobility by reducing unnecessary waiting times (Lee et al., 2013) and promote pedestrian compliance with traffic signals (Van Houten et al., 2006). Moreover, accessible push buttons that incorporate audio-tactile signals may be especially beneficial to blind and vision-impaired pedestrians (Barlow and Franck, 2005). In the advent of autonomous driving, the pedestrian push button remains an effective solution to mediate conflicts and improve pedestrian safety; however, the system may not be available at every intersection and midblock location. Furthermore, pedestrians tend to cross at convenient locations that present shorter delays (Ravishankar and Nair, 2018). Therefore, we followed an iterative design process to devise a concept where pedestrians can utilize the AR glasses to negotiate a crossing opportunity with approaching AVs. Prototypes of varying fidelities were created and improved through internal discussions among the authors. Additionally, two pilot studies (with a total of four participants) were conducted prior to the main investigation. User interactions were modeled after those used with the pedestrian button, comprising three stages, as illustrated in **Figure 1** and described in greater detail as follows.

Sending a crossing request: As a safety prerequisite, predicting pedestrian crossing intentions based on parameters such as the pedestrian dynamics, physical surroundings, and contextual scene information is one of the most critical tasks of AVs (Ridel et al., 2018). However, many challenges remain to be overcome in achieving a reliable and robust solution. For this reason, we implemented a user-initiated communication approach with pedestrians explicitly indicating their crossing intents for a greater sense of control. Users can send a crossing request to all nearby AVs by quickly tapping a touch surface on the temples of AR glasses. While various methods for controlling the AR glasses exist, the tapping gesture was selected for its simplicity and ease of prototyping. Additionally, it is widely employed in wireless earphones, smart glasses, and smart eyewear (e.g., Ray-Ban Stories).

Waiting for crossing signals: Analogous to how the pedestrian push buttons offer visual and audible feedback when pressed, the AR glasses displayed a text prompt acknowledging the crossing request. According to media reports on push-button usage in the United States, many people are unsure if the system is of value and even regard them as placebo buttons, with their presence only offering an "illusion of control" (Prisco, 2018). The confusion has arisen mainly because the push buttons are inoperative during off-peak hours or have been supplanted by more advanced systems (e.g., traffic sensors) and kept only for accessible features (Prisco, 2018). Considering these user frustrations stemming from a lack of understanding regarding how a system works, we ensured that the text prompt briefly explains the workings of the AR glasses.

Receiving crossing signals: We developed three communication approaches to visually convey AVs' responses to the crossing request. The first approach involves placing a visual cue on each vehicle ("distributed response;") specifically, the AR glasses

render a green overlay that covers a vehicle's surface to indicate a yielding intent. The idea of an overlay is based on the futuristic digital paints that may be incorporated in automobiles by 2050 (AutoTrader, 2020). Given the lack of consensus regarding the optimal placement of visual cues on a vehicle's body, an overlay offers the advantage of being noticeable and visible from various angles. Green was chosen as the color to indicate "go" because of its intuitiveness (Dey et al., 2020b); we also assumed that possible confusion in perspectives (Bazilinskyy et al., 2019) is less likely to occur when the user initiates the communication. The second approach entails the use of a single visual cue, in this case, an animated forward-moving pedestrian crossing, to convey the intentions of all cars ("aggregated response.") The zebra crossing is a widely recognized traffic symbol that numerous eHMI studies have investigated (Löcken et al., 2019; Nguyen et al., 2019; Dey et al., 2021; Prattico` et al., 2021; Tabone et al., 2021b); its forward movement indicates the crossing direction (Nguyen et al., 2019), and the markings have high visibility (Löcken et al., 2019). The third approach combines both types of visual cues by displaying car overlays and an animated zebra crossing simultaneously. This approach was implemented based on study findings from Hesenius et al. (2018), taking into account the possibility of participants having different preferences regarding different combinations of cues.
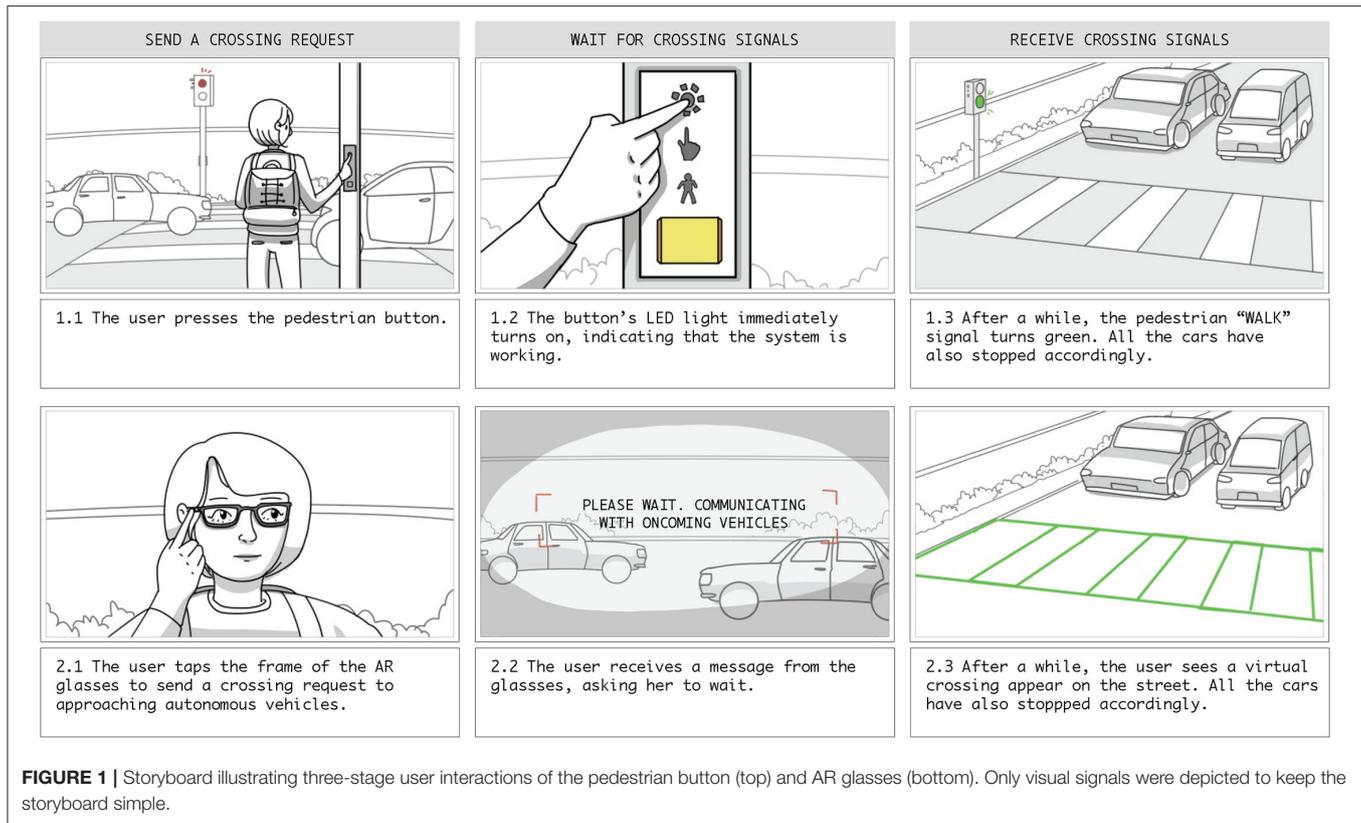
4. EVALUATION STUDY

4.1. Study Design

Given that the AR eHMI concepts were designed for a multi-vehicle traffic situation, a comparison to a currently implemented system would yield relevant insights into pedestrian preferences and crossing experiences. Therefore, we decided on a within-subjects study design with four experimental conditions: a baseline pedestrian push button and three wearable AR concepts with different communication approaches—aggregated response (AR crosswalk), distributed response (AR overlay), and both the aforementioned types (refer to **Figure 2**). To minimize carryover effects, we changed the order of presenting the concepts from one participant to another using a balanced Latin Square. We kept factors that might influence pedestrian behavior, such as vehicle speed, deceleration rates, and gaps between vehicles the same across all conditions. The participants' experimental task was to stand on the sidewalk, several steps away from traffic, and cross the street with the assistance of a given design concept.

4.2. Participants

To determine the required sample size, an a priori power analysis was performed using G*Power (Faul et al., 2009). With an alpha level of .05, a sample of 24 participants was adequate to detect a medium effect sizes (Pearson's r = 0.25) with a power of .81 (Cohen, 2013) for our measures.

We recruited 24 participants (62.5% female; 18–34 age range) through social media networks and word of mouth. The participants included working professionals and university students who had been living in the current city for at least 1 year and who could speak English fluently. All participants were required to have normal or corrected-to-normal eyesight,

FIGURE 1 | Storyboard illustrating three-stage user interactions of the pedestrian button (top) and AR glasses (bottom). Only visual signals were depicted to keep the storyboard simple.

as well as no mobility impairment. Of our participants, 13 had tried VR a few times, and three had extensive experience with it. Meanwhile, only two participants reported having experienced AR once. Thirteen participants required prescription glasses; the remaining 11 had normal visual acuity, three of which had undergone laser eye surgery, and one was using orthokeratology (i.e., corneal reshaping therapy) to correct their vision. The study was conducted at a shared workspace in Ho Chi Minh City (Vietnam), following the ethical approval granted by the University of Sydney (ID 2020/779). Participants in this study did not receive any compensation.

4.3. VR Prototype

Apparatus. The VR prototype was developed using the Unity[2] game engine and experienced with the Oculus Quest 2 VR system[3]. The head-mounted display (HMD) provides a fully untethered 6DOF experience and hand tracking feature, allowing users to walk around freely and engage in VR naturally with their hands (**Figure 3**). The experiment was conducted in an 8x3-meter open floor space, where participants were able to physically walk two-thirds of the street before being teleported to the other side. The (auto) teleportation was used to overcome HMD tracking space limits and to ensure that participants could observe how the visual cues disappeared and the AVs resumed driving after their crossing.

Virtual environment. The virtual environment was modeled using commercially available off-the-shelf assets. The scene featured an unmarked midblock location on a two-way urban street. Pedestrian crossing facilities, including traffic lights and zebra crossings, were only available under the experimental condition where pedestrians crossed the street using the pedestrian push button. To create a more realistic social atmosphere, we used Mixamo 3D characters[4] to replicate human activities on the sidewalk: some individuals were exercising while others were speaking with one another. Additionally, an urban soundscape with bird chirping sounds and traffic noise was included.

The vehicles used in this experiment were obtained from the Unity Asset Store and comprised a black/orange sports car, a silver sedan, and a white hatchback to create a more natural perception of traffic. Despite their model differences, these vehicles had similar sizes and kinematic characteristics, both of which were found to influence pedestrian experience and behavior Dey et al. (2017). Vehicular traffic was composed of fully automated cars (Level 5) (SAE, 2021) traveling in both lanes. To create the perception of autonomous driving, we did not model people inside and implemented a futuristic Audi e-tron sound[5] for each vehicle. The number of vehicles in each lane varied, but they consistently traveled with impassable gaps to ensure that participants could not cross without the AVs yielding. In

---

[2]https://unity.com/
[3]https://www.oculus.com/quest-2/

[4]https://www.mixamo.com/
[5]https://www.e-tron-gt.audi/en/e-sound-13626

FIGURE 2 | Simulation environment and interfaces included in the evaluation: **(A)** Pedestrian push button; **(B)** AR crosswalk; **(C)** AR overlay; **(D)** AR-combined.

the simulation, the vehicles were spawned at a location hidden from the participants' view; they started accelerating and driving at approximately 30 km/h before making a right turn. When responding to pedestrians' crossing requests, the vehicles slowed down at a distance of 19 m, following the safe stopping distance recommended in urban zones[6]. They came to a complete stop at 1.5 m from the designated crossing area and only resumed driving once the participants had reached the other side of the road (refer to **Figure 4**).

Evaluated concepts. We commissioned a game artist to create a 3D model of the Prisma TS-903 button[7] used in the city where the study took place. In VR, participants could use their hands to engage with the button in the same way they would in real life (refer to **Figure 2A**). To experience the wearable AR concepts, participants used the VR headset as if it was a pair of AR glasses and were instructed to tap on its side whenever they planned to cross. On Oculus Quest 2, this type of tap gesture was not available; it was prototyped by creating an invisible collision zone around the HMD that detects any contact with the user's fingertips. The tapping immediately triggers sound feedback and displays an HUD text prompt "*Please wait. Communicating with*
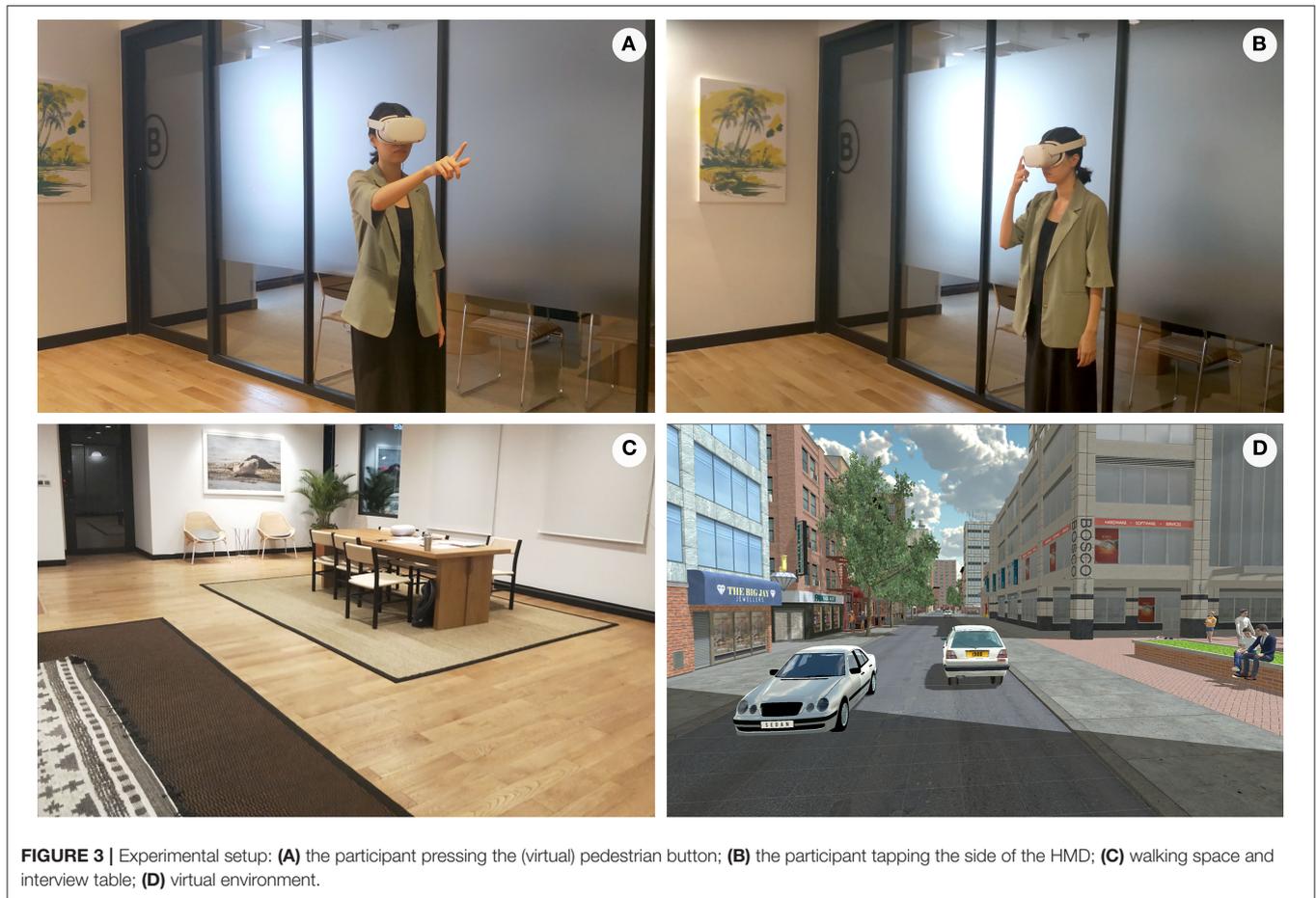
oncoming vehicles" in users' primary field of vision. After 9 s, all AVs responded to the pedestrian crossing request by decelerating at a distance of 19 m and displaying the car overlays. However, those with a short stopping distance (already approaching the pedestrians when the request was received) continued to drive past to avoid harsh braking. To account for these cars, a 3-s delay was put in place to make sure that the AR zebra crossing only appeared when the crossing area was safe. The design of the zebra crossing was inspired by the Mercedes-Benz F 015 concept[8], with bright neon green lines and flowing animation (refer to **Figure 2B**). The car overlay was made of semi-transparent emissive green texture and appeared to be a separate layer from the car (refer to **Figure 2C**). Both the zebra crossing and the car overlay are conformal AR graphics situated as parts of the real world. In addition to visual cues, we offered audible signals to indicate wait time (slow chirps) and crossing time (rapid tick-tock-tick-tock). These sounds are part of the Australian PB/5 push button signaling system[9], and they were implemented across four experimental conditions.

---

[6]https://roadsafety.transport.nsw.gov.au/speeding/index.html
[7]https://www.prismatibro.se/en/prisma-ts-903-eng/

[8]https://www.mercedes-benz.com/en/innovation/autonomous/research-vehicle-f-015-luxury-in-motion/
[9]https://www.maas.museum/inside-the-collection/2010/04/16/pedestrian-button-1980s-australian-product-design-pt2/

FIGURE 3 | Experimental setup: **(A)** the participant pressing the (virtual) pedestrian button; **(B)** the participant tapping the side of the HMD; **(C)** walking space and interview table; **(D)** virtual environment.

FIGURE 4 | A top-down view of the virtual environment zooms in on the midblock location where participants made the crossing. Dotted blue arrows indicate the travel direction of AVs. Solid blue arrows indicate where the AVs (on each lane) begin to decelerate and where they come to a complete halt. The white circle indicates the pedestrian position at the start of each experimental condition.

4.4. Procedures

After the participants had signed up for the study, a screening questionnaire was used to obtain their demographic information, including age group, gender, English proficiency level, occupation, nationality, length of stay in the current city, walking issues, and eye conditions. On the day of the study, we welcomed the participants and gave them a brief overview of the study and the related tasks. The participants were then asked to read and sign a consent form. Following a quick introduction to the VR system, we asked the participants to put on the HMD and adjust it until they felt comfortable and could see the virtual environment clearly. A glasses spacer was inserted in the HMD

such that the participants could wear the headset with their glasses on.

Before beginning the experiment, the participants took part in a familiarization session in which they practiced crossing the street and interacting with virtual objects with their hands. Prior to each experimental condition, we presented the participants with an image of the pedestrian push button or the AR glasses to gauge their familiarity with the technology and inform them about the system with which they would be engaging. However, they were not made aware of the differences between the wearable AR concepts. After each condition, the participants removed their headsets and completed a series of standardized questionnaires at a nearby table. We also ensured that no participant was experiencing motion sickness and that all could continue with the experiment. After all the conditions had been completed, we conducted a semi-structured interview to gain insights into their experiences.

4.5. Data Collection

After each experimental condition, we monitored participants' simulator sickness with the single-item Misery Scale (Bos et al., 2010). If the rating was four or higher, the study would be suspended. We then measured perceived cognitive load with the NASA Task Load Index (NASA-TLX) (Hart and Staveland, 1988) on a 20-point scale. The questionnaire has six workload-related dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. These dimensions were combined into one general cognitive load scale Cronbach's $\alpha = 0.783$). To assess trust in human-machine systems, we used a 12-item trust scale (Jian et al., 2000). The first five items provided an overall distrust score (Cronbach's $\alpha = 0.896$); the next seven items provided an overall trust score (Cronbach's $\alpha = 0.941$). Finally, the 10-item System Usability Scale (Brooke et al., 1996) was used to measure usability (Cronbach's $\alpha = 0.909$). All the questionnaires were explained to the participants and administered under supervision. We also instructed participants to assess the prototyped systems instead of the VR representation.

After the completion of all experimental conditions, the participants were asked to rank the systems from 1 (most preferred) to 4 (least preferred). Additionally, a semi-structured interview was conducted to gain a better understanding of their overall experience, the reasoning behind their preferences, and their perspectives on various system aspects and the VR simulation.

4.6. Data Analysis

Questionnaires: We first calculated summary statistics and created data plots to investigate the data sets. We assessed the normality of data using Shapiro-Wilk tests and a visual inspection of their Q-Q plots. Because most data have non-normal distribution, we used the non-parametric Friedman test to determine any statistically significant differences in questionnaire outcomes. In case of significant differences, we performed Dunn-Bonferroni procedure for multiple pairwise comparisons as *post-hoc* tests. We considered an effect to be

significant if $p < .05$. IPM SPSS version 28 was used for all statistical analyses.

Interviews: Post-study interviews were transcribed by the interviewer with the assistance of an AI-based transcription tool. Two coders performed an inductive thematic analysis (Braun and Clarke, 2006) to identify and interpret patterns (themes) within the data. The first coder (TT) had extensive knowledge of the study, while the second coder (YW) was not involved in its conception and implementation. This approach enabled us to have a more complete and unbiased look at the data.

The analysis began with the first coder selecting a subset of six interviews (25% data units) with good representativeness. The first round of coding was performed independently by both coders, followed by a discussion to agree on the coding frame. In the second round, the first coder applied the coding frame to all interviews. Finally, we examined the themes and patterns that emerged, which composed part of the Results section.

5. RESULTS
5.1. Concept Ranking

Regarding the top preference (first ranking), approximately half of the participants preferred the AR concept incorporating both the animated crosswalk and car overlays, while one-third favored the pedestrian button. The *AR overlay*, followed by the *AR crosswalk*, was the least preferred (refer to **Table 1**).

A Friedman test showed a significant difference in the mean rankings among concepts ($\chi^2(3) = 29.850, p < 0.001$). *Post-hoc* tests revealed that the *AR-combined* ($mdn = 1.0$) was rated significantly higher than the *AR crosswalk* ($mdn = 3.0$) ($z = 1.375, p_{corrected} = 0.001$) and *AR overlay* ($mdn = 4.0$) ($z = 1.875, p_{corrected} = 0.000$) but not the pedestrian button ($p_{corrected} = 0.705$). The pedestrian button ($mdn = 2.0$) was rated significantly higher than the *AR overlay* ($mdn = 4.0$) ($z = -1.292, p_{corrected} = 0.003$).

5.2. SUS

Based on the grade rankings created by Bangor et al. (2009), the System Usability Scale (SUS) scores of the *Button* and the *AR-combined* were considered as "excellent." The *AR crosswalk* and the *AR overlay* had lower scores which were in the "good" range (refer to **Table 2**). A Friedman test indicated a significant main effect of the concepts on the usability scores ($\chi^2(3) = 10.808, p = 0.013$). *Post-hoc* analysis revealed that the usability scores were statistically significantly different between the *Button* ($mdn = 85$) and the *AR overlay* ($mdn = 77.50$), ($z = 1.063, p_{corrected} = 0.026$), as shown in **Figure 5**.

5.3. NASA-TLX

Descriptive data analysis showed that the overall scores were low for all concepts; however, the *AR-combined* elicited the least cognitive load (refer to **Table 2**). A Friedman test showed a significant difference in the overall mean scores ($\chi^2(3) = 11.535, p = 0.009$). *Post-hoc* tests revealed that the *AR crosswalk* received significantly higher cognitive load scores ($mdn = 18.33$) compared to the *AR-combined* ($mdn = 10.84$) ($z = 1.000, p_{corrected} = 0.044$), as shown in **Figure 6**.

TABLE 1 | Ranking results by frequency of nomination.

	Button	AR crosswalk	AR overlay	AR-combined
1st rank	8	2	1	13
2nd rank	8	6	1	9
3rd rank	5	8	9	2
4th rank	3	8	13	0

TABLE 2 | Mean (M) and standard deviation (SD) for NASA-TLX scores, SUS scores, and Trust Scale ratings.

	Button	AR crosswalk	AR overlay	AR-combined
	(M / SD)	(M / SD)	(M / SD)	(M / SD)
SUS	**85.31** / 12.30	**77.29** / 19.78	**74.48** / 20.61	**85.10** / 13.58
NASA-TLX	**20.94** / 13.87	**20.90** / 14.55	**21.35** / 14.23	**13.99** / 8.53
Trust (subscale)	**5.83** / .96	**5.02** / 1.39	**4.82** / 1.39	**5.52** / 1.17
Distrust (subscale)	**1.85** / 1.10	**2.32** / 1.21	**2.60** / 1.45	**1.99** / .94

Regarding subscales, the Friedman test found a statistically significant effect of the concepts on *temporal demand* ($\chi^2(3) = 12.426, p = 0.006$) and *frustration* ($\chi^2(3) = 8.392, p = 0.039$). The *post-hoc* tests showed no significant differences ($p_{corrected} > .05$). However, the uncorrected p-values indicated that the *AR-combined* received significantly lower scores in *temporal demand* compared to all other concepts.

5.4. Trust Scale

According to descriptive data analysis (refer to **Table 2**), the participant's trust in the three AR concepts was lower than in the *Button*, with the lowest trust in the *AR overlay*. Results from a Friedman test found a significant difference in the mean scores of trust ratings ($\chi^2(3) = 14.724, p = 0.002$). *Post-hoc* tests revealed that the *Button* ($mdn = 6.00$) received significantly higher trust ratings compared to the *AR crosswalk* ($mdn = 5.07$) ($z = 1.021, p_{corrected} = 0.037$) and the *AR overlay* ($mdn = 4.79$) ($z = 1.188, p_{corrected} = 0.009$), as shown in **Figure 7** on the left.

Participants' distrust in the three AR concepts, conversely, was higher than that in the *Button*, with the strongest level of distrust being shown in the *AR overlay* (refer to **Table 2**). A Friedman's test showed a significant difference in the mean ratings ($\chi^2(3) = 15.556, p = 0.001$). *Post-hoc* tests revealed that the *AR overlay* ($mdn = 2.10$) received significantly higher distrust ratings compared to the *Button* ($mdn = 1.50$) ($z = -1.167, p_{corrected} = 0.010$), as shown in **Figure 7** on the right.

5.5. Qualitative Feedback

This section presents the primary themes that emerged from our qualitative data analysis, providing insight into the participants' perceptions of wearable AR concepts and the design features that influenced their experiences.
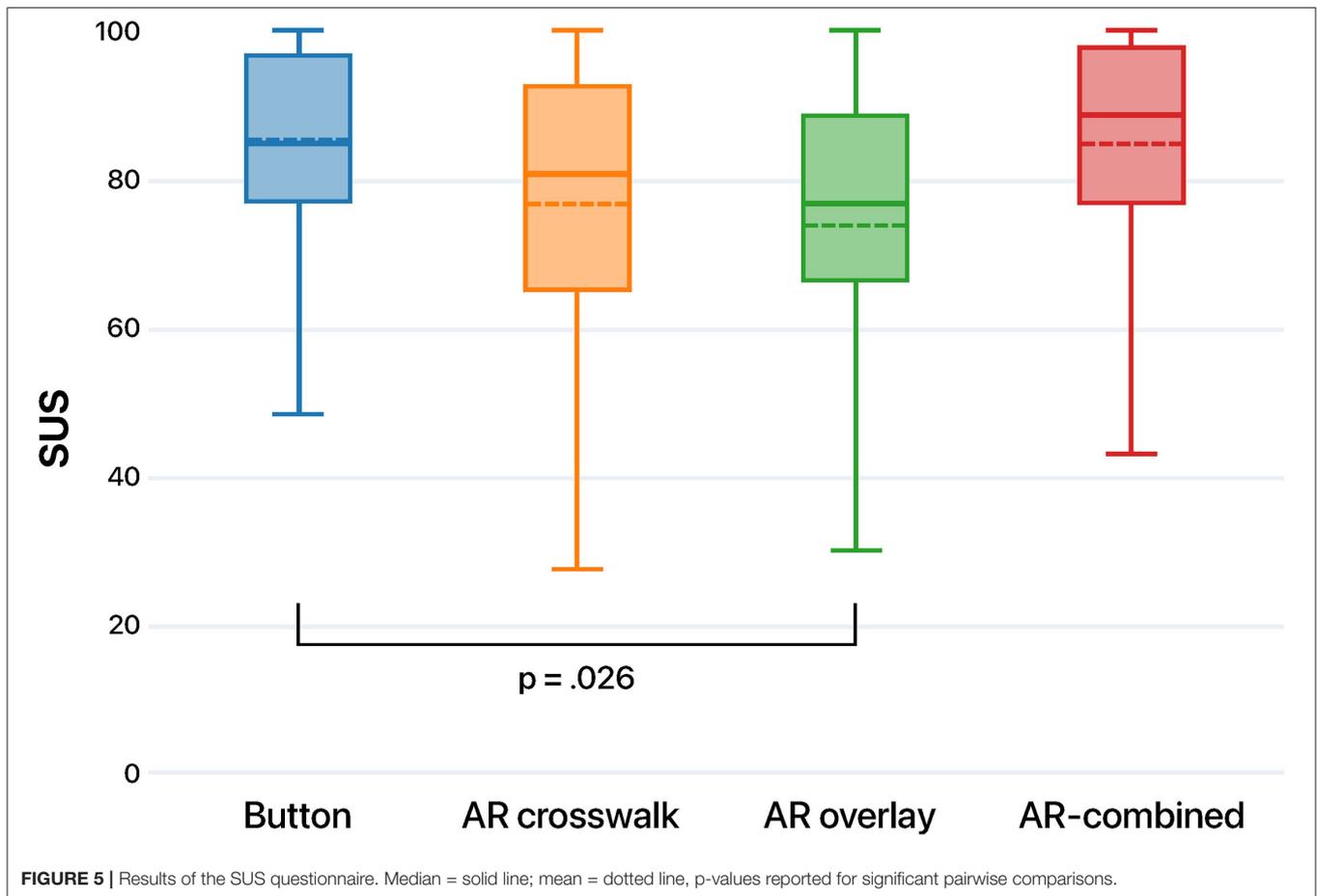
(1) Wearable AR concepts were unfamiliar yet exciting: The post-study interviews showed that the participant's familiarity with the design solutions appeared to influence their trust in them. The pedestrian push button was perceived as highly familiar by a noticeable ratio of the participants ($n = 10$). This sense of familiarity was often linked to past experiences ($n = 8$) and had frequently resulted in feelings of confidence while crossing ($n = 7$). For example, P23 stated, *"I feel safer because it's something that I'm used to. I have the feeling that it's guaranteed."* Wearable AR applications, on the other hand, were regarded as novel and less familiar than the traditional infrastructure ($n = 6$), which might hinder their uptake, especially in the older generation (P4 and P20). As a result, several participants recommended that providing onboarding tutorials (P7) or a user manual (P2) might benefit their adoption. Furthermore, a number of participants stated that additional exposure to wearable AR applications is necessary to establish their dependability ($n = 7$) - *"I have only experienced it once. Maybe I need to interact and use it a few times. I need to try it more to know if it's reliable"* (P21). P7 added that knowledge of relevant statistics, such as the number of users of the AR system, could also contribute to an increase in trust.

Despite the unfamiliarity, wearable AR solutions were frequently described as exciting and cool (n = 5). As commented by P2, *"It's like I have mind control and being able to stop all the cars."* In contrast, the pedestrian push button was deemed to be a conventional system to support pedestrian crossing ($n = 4$), referred to as *"very old school"* (P16) and *"less technologically advanced"* (P13).

(2) Wearable AR offered both advantages and disadvantages as a personal device: The pedestrian push button baseline enabled a direct comparison of a solution based on personal devices with an infrastructure-based solution, producing a variety of insightful perspectives from the participants on the personal nature of AR eHMIs. The analysis showed that one of the most commonly noted advantages of wearable AR concepts is the increased flexibility of crossing locations ($n = 4$), as opposed to the fixed installation of the pedestrian push buttons. P7 found it particularly useful when *"[she] wants to cross the street in a hurry"* (P7). Furthermore, P5 noted that the precision of requests sent to the vehicles could result in higher efficiency— *"normal vehicles usually focus on the street; maybe they will miss my request to cross the street. If I use the AR glasses, it'd be quicker I think."* Nonetheless, cost ($n = 4$) and data privacy ($n = 4$) were identified as two of the most significant barriers to personal devices being adopted over public infrastructure. Two participants also raised concerns about circumstances where they might forget the personal device at home (P6 and P9) or do not wish to wear the AR glasses at times (P9). Similarly, personal devices were perceived to be inferior to public infrastructure in terms of liability ($n = 3$) and maintenance ($n = 2$). As commented by P8, *"because the button is of the government, if there's something happened, we can find somebody to blame,"* while P4 stated that *"[she] believe[s] there will be someone taking good care of a public system."*

(3) The physical form factor of the AR glasses was found to influence their acceptance: The idea of wearing glasses (or even contact lenses) was not appealing to individuals who had undergone eye corrective surgery. The concern was less about
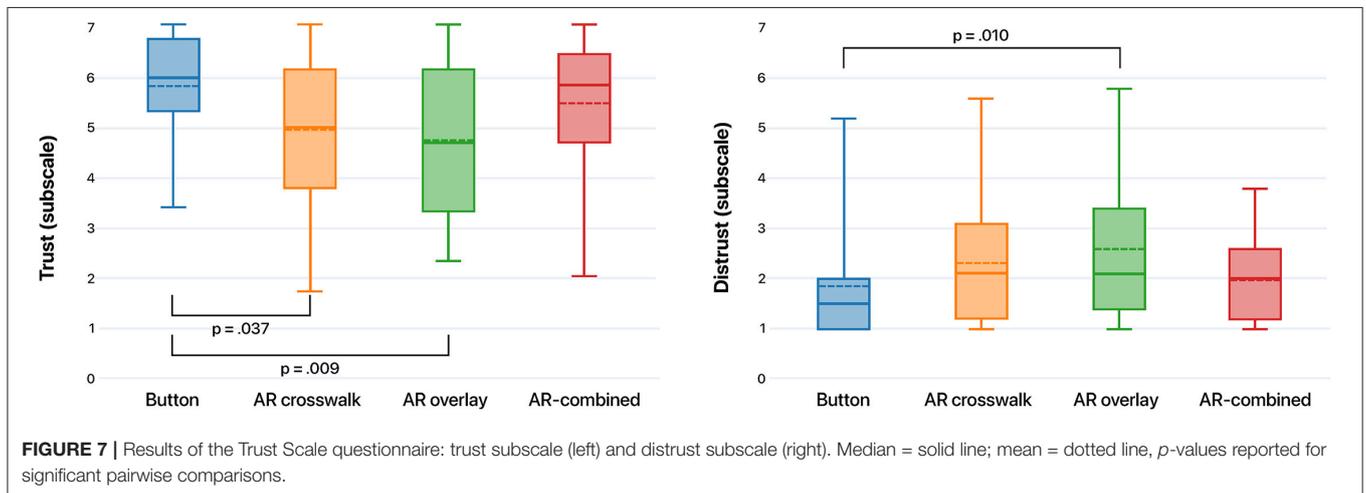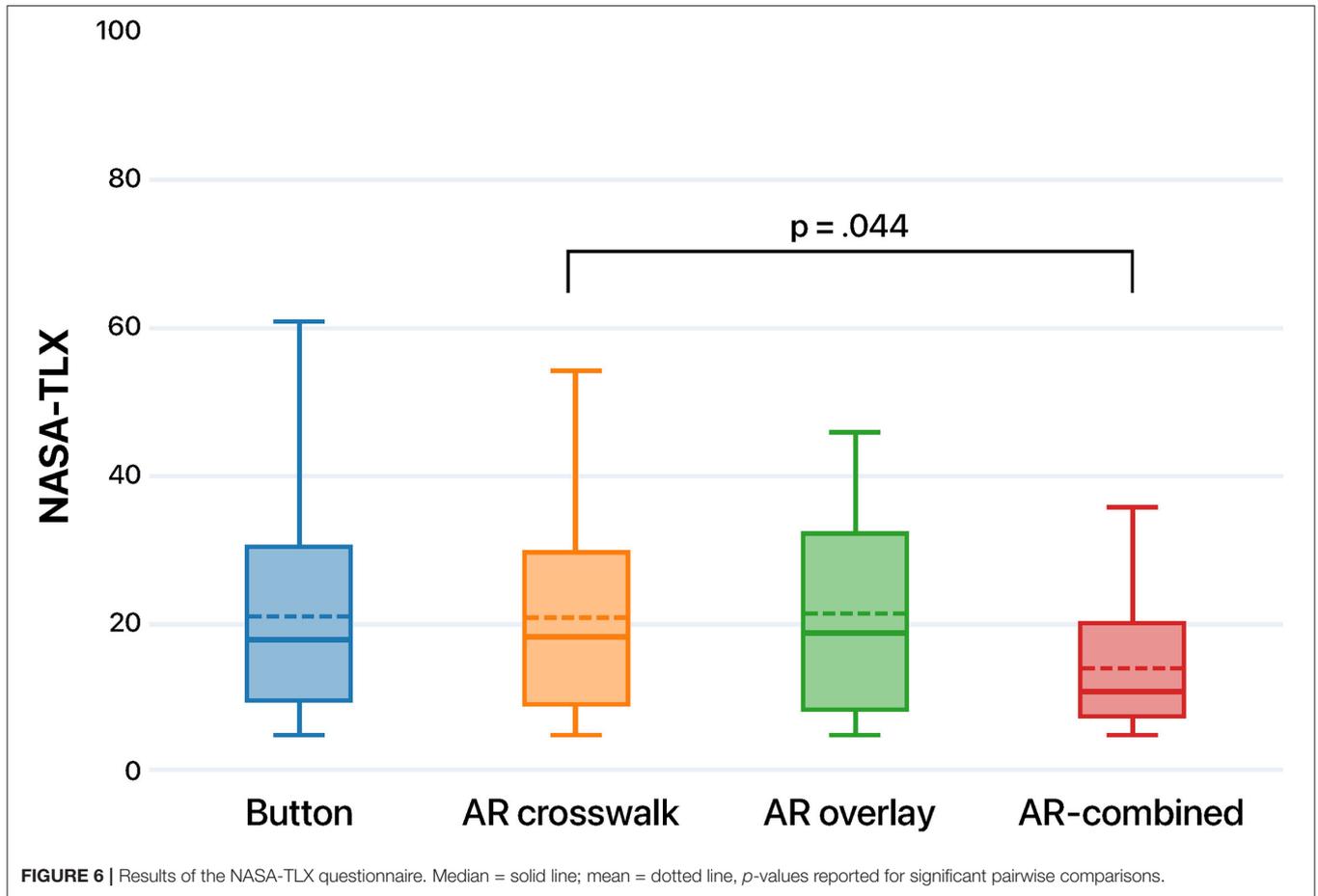
FIGURE 5 | Results of the SUS questionnaire. Median = solid line; mean = dotted line, p-values reported for significant pairwise comparisons.

the aesthetic qualities of the AR glasses but more about the (re)dependence on eyewear on a daily basis (*n* = 3). Three participants questioned the necessity of using AR glasses to aid in crossing. Furthermore, two people suggested smartphones (P1), smartwatches (P8), and AVs' pedestrian detection feature (P8) as alternative solutions to wearing AR glasses. Nonetheless, three participants identified the potential of using AR glasses for multiple purposes, such as reading the news and watching television (P9), rather than solely assisting in crossing.

(4) **User-initiated communication provided a sense of control**: When questioned about the preferred mode of interaction, all participants (*n* = 24) responded that they favored sending a crossing request to AVs rather than waiting for the vehicles to detect their intentions. We found that the participants' reasoning regarding this preference revolved around two aspects. First, some participants were skeptical about the ability of AVs to capture intricate human intentions (*n* = 13), stating that pedestrians may cross the street "*spontaneously*" (P12), "*change their minds*" quickly, or move in ways that suggest something unintentionally (P16). One participant doubted the reliability of algorithms that learn from "*previously fed data*" (P13). Second, some participants preferred to have some control over the interaction (n = 9); according to them, proactive communication with AVs was deemed critical for ensuring accuracy and hence

safety (n = 9). Concerns about the passivity and uncertainty associated with waiting were also mentioned: "*I have no way to know that whether they will stop or not. What if [the cars] just keep moving?*" (P21).

The qualitative analysis further suggested that the participants preferred a digital approach over bodily gestures (e.g., waving hands), owing to the lack of confidence that AVs would all be able to observe their signals (*n* = 7). For example, P15 stated, "*If I raise my hand, I'm not sure if all cars see it.*" Nonetheless, whereas the integration with traffic lights enables the pedestrian buttons to operate effectively in mixed traffic situations, the practicality of wearable AR to communicate with manual vehicles (*n* = 5) and the extent to which human drivers cooperate (*n* = 6) were questioned. Furthermore, six participants expressed reservations about potential traffic disruptions in the event of many road users using the AR glasses for street crossing. P14 stated, "*what if there were 10, 20 people wearing glasses, but they do not cross the street at the same time?*".

(5) **Clear communication mechanisms with AVs influenced the perceived safety**: We found that the perceived connection between the system used and AVs influenced the participants' feeling of safety. Regarding the AR glasses, the connection was seen as direct and explicit (*n* = 7). The provision of visual cues

FIGURE 6 | Results of the NASA-TLX questionnaire. Median = solid line; mean = dotted line, *p*-values reported for significant pairwise comparisons.

FIGURE 7 | Results of the Trust Scale questionnaire: trust subscale (left) and distrust subscale (right). Median = solid line; mean = dotted line, *p*-values reported for significant pairwise comparisons.

assured the participants that the connection was *"established"* and would continue to be maintained during their crossing, as reported by P1: *"I assumed that the vehicles would be waiting for me to finish the crossing. They will allow me as much as possible time to cross [...] because they may be connected to my glasses and aware of my presence."* In the case of the pedestrian button, user feedback revealed divided viewpoints. Eight participants

were puzzled as to how the system "talked" to the vehicles. P21 thought that *"the digital context [was] missing*, while P17 viewed the two entities as *'disconnected.'"* Meanwhile, nine participants contended that the AVs came to a halt due to a changing traffic signal. P18 highlighted that the vehicles *"might have a sensor to read the color [sic] of the traffic light."* It was this interpretation and confidence in the ability of the traffic lights to regulate

traffic that allowed these participants to feel more at ease in the interaction than the other group.

It is worth noting that several participants paid close attention to the technical aspects of the connection, highlighting possible risks that might occur with wearable AR concepts (n = 10). Five participants voiced concerns about the potential malfunctions of individual entities, which can imperil the operation of the integrated system. P6, e.g., mentioned a scenario where *"one vehicle does not comprehend the signal."* Three participants suggested connection failures, such as internet disconnections (P11) and signal transmission delays (P14). Two participants highlighted that the system might suffer from malicious manipulation (e.g., hacking).

(6) The combined approach provided extra security: Several participants reported that seeing the zebra crossing and car overlays simultaneously boosted their confidence (n = 12). In this regard, they reasoned that the dual cues provided *"extra"* security by exhibiting a strong integration of various entities (i.e., the AR glasses and the vehicles). As P19 explained, *"If there is a misconfiguration or anything that is not synchronized, I may be aware of that and know when the system has an issue."*

Furthermore, we noted several remarks on the perceived usefulness of each visual cue, shedding light on why their presence was instrumental in the pedestrian crossing experience. Approximately half of the participants interpreted the car overlay as a direct response from each vehicle to their crossing request (n = 11). P1, for instance, felt as though *"[the vehicles were] actually listening"* and that the connection worked. In scenarios lacking these individual confirmations, participants reported feeling uncertain about the AV yielding behavior (n = 3). As P21 expressed, *"[...] what if there are three or four lanes of cars? If I don't see this green thing, I feel a little bit worried. Maybe some cars will stop, and some will not stop,"* With respect to the zebra crossing, a sizeable proportion of the participants regarded it as a clear crossing signal due to its high visibility (n = 3) and familiarity (n = 9). The AR marking superimposed on the street also served as a visual cue indicating where the AVs would stop (n = 5).

6. DISCUSSION

In this section, we discuss the findings in relation to our research questions and reflect on the limitations of our study.

6.1. Preference for Wearable AR Concepts (RQ1)

The quantitative results indicated that employing wearable AR to aid AV-pedestrian interaction was a viable approach. This is evident in the case of the *AR-combined* concept, which was ranked higher than the baseline pedestrian button and significantly reduced the street-crossing cognitive load. In addition, even though the concept was rated marginally lower in usability and trust due to its unfamiliar nature, no statistically significant differences could be found. However, not all the wearable AR concepts performed similarly. The *AR overlay* and *AR crosswalk* both significantly induced higher

distrust and lower trust compared to the baseline; they also received lower usability scores. The discrepancy in the ratings among the wearable AR concepts leads us to infer that the communication approach employed strongly influenced the pedestrians' subjective experiences. The qualitative feedback confirmed this observation and further suggested that the extent to which pedestrians preferred to use AR glasses to interact with AVs was also influenced by their perception of wearable AR technology.

With respect to wearable AR technology, the semi-structured interviews revealed important factors influencing pedestrians' adoption of AR solutions in interacting with AVs, including costs, data privacy, technical defects, liability risks, maintenance duties, and form factors. Although these problems were not widely discussed among the participants, they reinforce expert opinions that wearable AR should not be the sole means for pedestrians to cross the street or engage with AVs in general (Tabone et al., 2021a). Several participants suggested alternative methods of communication with AVs, such as using smartphones, which indicated that a user-initiated communication concept was appreciated more than the underlying AR technology. This inclination might be explained by smartphones' present ubiquity and their ecosystem of applications. As wearable AR is becoming more pervasive with continuous AR experiences (Grubert et al., 2016)—e.g., a pedestrian may use wearable AR for navigational instructions, communication with AVs when crossing the road, or retrieving information about the next train home—we hypothesize that pedestrian attitudes may shift in the future.

Concerning interactions with AVs in safety-critical settings, the participants unanimously agreed on the need to make their crossing intentions known to AVs. This finding is consistent with a prior study on bidirectional communication between pedestrians and AVs (Colley et al., 2021; Epke et al., 2021), which showed that a combination of hand gestures and receptive eHMIs was the most desired method of communication. However, while hand gestures have been previously observed to have limitations in terms of false-positive (Epke et al., 2021) or false-negative detection (Gruenefeld et al., 2019), a digital approach was viewed as safer and more trustworthy in our study. Additionally, using wearable AR for bidirectional communication not only ensures that AVs accurately interpret pedestrian intentions but might also eliminate potential confusion about AV non-yielding behaviors (Epke et al., 2021). For example, AR may be utilized to increase system transparency by explaining long wait times or a refusal to yield. According to prior study, explanations of AI system behavior can promote trust in and acceptance of autonomous driving (Koo et al., 2015). A substantial body of literature on explainable AI has focused on drivers' perspectives; nevertheless, a survey article has argued that the provision of meaningful explanations from AVs could also benefit other stakeholders (e.g., pedestrians) (Omeiza et al., 2021).

It was anticipated that wearable AR could readily enable targeted and high-resolution communication between AVs and individual pedestrians; a user could be assured that the AVs were addressing them because the device was used individually. In our study, the clarity of recipient was further reinforced when

pedestrians were the ones who initiated the communication. However, despite the advantages of wearable AR concepts in delivering unambiguous messages, we found that the aspect of individual perceptions merits further discussion. According to qualitative data, the participants were concerned whether the proposed AR solutions would benefit urban traffic as a whole, as revealed by the raised concerns about frequent crossing requests. In this regard, P17 made a noteworthy comment about a possible shared perception among wearable AR users with regard to visual signals: *"if there are also other people, then I will prefer the crosswalk. People will be crossing the street at the same time and in the same place."* This comment leads us to believe that in certain situations, a shared AR experience (Rekimoto, 1996), where multiple users can see the same virtual elements, may help guide pedestrian traffic more efficiently. As a result, personal and shared (augmented) reality should both be considered when designing AR eHMIs.

6.2. Communication Strategies (RQ2)

The quantitative findings suggested that the *AR-combined* concept performed better than the *AR crosswalk* (aggregated response) and *AR overlay* (distributed response) across all measures, despite a statistically significant difference only being observed in the concept ranking. In terms of cognitive load, the uncorrected p-values indicated that combining visual signals could considerably reduce pedestrians' temporal demand as compared to presenting each cue individually, which might mean that when using the *AR-combined* concept, the participants felt less time-pressured as they crossed the road. This tendency was supported by qualitative findings where the participants reported feeling more confident during their crossings. To further understand the benefits and drawbacks of each communication approach, as well as why they were able to complement one another, we discuss them in further detail as follows.

(1) **Aggregated response**: As one of the most widely recognized traffic symbols, the marked pedestrian crossing was chosen to show an aggregated response from all incoming vehicles, indicating that they were aware of the pedestrians and would yield to them. We expected that this communication strategy would reduce the amount of time and effort required to read implicit or explicit cues from many vehicles. However, the analysis revealed that while the crosswalk indicated a clear signal to cross and a designated crossing area, the participants remained unsure of the AVs' yielding intention and relied more on vehicle kinematics to make crossing decisions. This finding appears to contradict those of Löcken et al. (2019), in which the participants began crossing as soon as the smart road's crosswalk lights had turned green, without waiting for AV signals. We believe that the difference in traffic scenarios (one vehicle vs. multiple vehicles) and the underlying technologies (smart infrastructure vs. personal devices) between the two studies might have contributed to divergent outcomes.

Notably, user interviews indicated that the participants were not familiar with the notion of connected vehicles; their hesitation persisted even after the leading vehicles had come to a complete stop. The fact that pedestrians do not perceive all AVs as a single system has also been observed during an evaluation of the

"omniscient narrator," where one representative vehicle was in charge of aural communications (Colley et al., 2020a). However, it is worth noting that in our study, the crossing signal originated from the AR glasses rather than from one of the AVs. Therefore, it would be useful to further investigate the difference between the two approaches. Moreover, while (Colley et al., 2020a) expressed reservations about the practicality of aggregated communication in mixed traffic scenarios, we believe that the approach may be feasible with the introduction of connected vehicle technologies. Through the use of in-vehicle or aftermarket devices, vehicles of varying levels of automation can exchange data with other vehicles (V2V), roadside infrastructures (V2I), and networks (V2N) (Boban et al., 2018). Such connections may result in a gradual shift of pedestrian trust away from specific entities and toward the traffic system as a whole. For example, when responding to the Trust Scale questionnaire, several participants stated that they viewed AR glasses and AVs as a unified system.

(2) **Distributed response**: The multi-vehicle traffic situation highlighted the necessity for pedestrians to be guaranteed successful communication with every AV, as evident in the positive user feedback on the car overlay. However, a confounding factor was present in the results when some individuals overlooked the overlay, believing that the cars were *"always green."* We attributed the cause of this issue to the simulation of AR in VR, where the contrast between the augmented graphics and the "real" environment was not as accurate as it should have been. Additionally, we believe that the participants' attention might have been scattered in a scenario involving multiple vehicles. For instance, P21 stated that he had to turn left and right to observe the two-way traffic and, therefore, failed to notice *"the changing colors."* This issue of split attention in complex traffic situations might also present difficulties for distance-dependent eHMIs (Dey et al., 2020a), the encoded states of which change with the distance-to-arrival, as pedestrians may not notice the entire sequence.

Regarding the display of individual car responses in complex mixed traffic situations, the study findings of Mahadevan et al. (2019) have suggested that this approach would enable pedestrians to assess each vehicle's awareness and intent and distinguish AVs from other vehicle types. Nonetheless, even standardized eHMI elements could be problematic since each car manufacturer might opt for slightly different designs. As a result, we believe that wearable AR may present a good opportunity for consistent visual communication across vehicles and serve as a clear indicator of their current operation mode (manual vs. autonomous) as needed.

6.3. Limitations and Future Work

First, the findings of our study drew on the experiences of a small number of university students and young professionals. Although we anticipate comparable outcomes, a larger representative sample would be beneficial, particularly in resolving some borderline quantitative results. Furthermore, past research indicates that cultural differences may cause eHMIs to not have the same favorable effect across countries (Weber et al., 2019). Given that the participants in our study largely came from the same cultural background (92% Vietnamese, 8% Indian) and had

similar habitual traffic behaviors, the feasibility of transferring the wearable AR concepts to differing cultures should be further investigated. Nevertheless, we argue that AR could easily offer personalized experiences, as opposed to vehicle-mounted or infrastructure-based eHMIs.

Second, the ecological validity of this study is limited by the use of a VR simulation. The virtual environment could not fully replicate the complex sensory stimuli found in the real world, and the safety associated with VR testing might have influenced individuals to engage in riskier crossing behaviors. Additionally, a few participants expressed anxiety over colliding with physical objects, despite our assurance otherwise. Nonetheless, the majority of the participants responded favorably to the simulation's realism, stating that they behaved similarly to how they would in the real world. They did not experience any particular motion sickness symptoms and were not affected by the short-distance teleportation implemented at two-thirds of their crossing. Existing literature also suggests that while achieving absolute validity and numerical predictions may not be possible, the VR method can effectively identify differences and patterns (Schneider and Bengler, 2020).

Finally, our study employed VR to prototype wearable AR concepts. Although this approach proved useful in overcoming the technical constraints of current AR HMDs, particularly in an outdoor setting, it was challenging for some participants to distinguish superimposed AR graphics from the virtual environment. To some extent, this issue confounded the results of the design concepts with car overlays (the *AR overlay* and the *AR-combined*), possibly causing them to be rated lower than they should have been. However, we believe that it did not invalidate the findings because the order of the four experimental conditions was counterbalanced, and the participants were able to recognize the visual cue in their second encounter. Furthermore, given the possibility of resolving this issue by contrasting display fidelity between AR and VR elements, we recommend that this VR simulation approach be considered in future study. With a large number of proposed AR design concepts in the literature, such as the nine prototypes created by Tabone et al. (2021b), comparison studies may provide intriguing insights into how AR systems best facilitate AV-pedestrian interaction.

7. CONCLUSION

This article has presented novel AR eHMIs designed to assist AV-pedestrian interaction in multi-vehicle traffic scenarios. Through a VR-based experiment, three wearable AR design concepts with differing communication approaches were evaluated against a pedestrian push button baseline. Our results showed that a wearable AR concept highlighting individual AV responses and offering a clear crossing signal is likely to reduce crossing pedestrians' cognitive load. Furthermore, enabling pedestrians to initiate the communication offered them a strong sense of control. This aspect of user control is currently underexplored in AV external communication research, pointing to important

future work in this domain. Finally, the adoption of wearable AR solutions depends on various factors, and it is critical to consider how VRUs without AR devices can interact with AVs safely and intuitively.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because The University of Sydney Human Research Ethics Committee (HREC) has not granted the authors permission to publish the study data. Requests to access the datasets should be directed to tram.tran@sydney.edu.au.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The University of Sydney Human Research Ethics Committee (HREC). The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

TT contributed to conceptualization, data collection, data analysis, and writing (original draft, review, and editing). CP and MT contributed to conceptualization, writing (review and editing). YW contributed to data analysis, and writing (original draft, review, and editing). All authors have read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomp.2022.866516/full#supplementary-material

Supplementary Video 1 | Video showcasing wearable AR prototypes and the experimental setting of the study.

REFERENCES

Woodhouse, A. (2020). *Cars of the future*. AutoTrader. Available online at: https://www.autotrader.co.uk/content/features/cars-of-the-future> (accessed March 17, 2022).

Azuma, R. T. (2019). The road to ubiquitous consumer augmented reality systems. *Hum. Behav. Emerg. Technol.* 1, 26–32. doi: 10.1002/hbe2.113

Bangor, A., Kortum, P., and Miller, J. (2009). Determining what individual sus scores mean: adding an adjective rating scale. *J. Usability Stud.* 4, 114–123. doi: 10.5555/2835587.2835589

Barlow, J. M., and Franck, L. (2005). Crossroads: Modern interactive intersections and accessible pedestrian signals. *J. Vis. Impairment Blindness* 99, 599–610. doi: 10.1177/0145482X0509901004

Bazilinskyy, P., Dodou, D., and De Winter, J. (2019). Survey on ehmi concepts: the effect of text, color, and perspective. *Transp. Res. Part F Traffic Psychol. Behav.* 67, 175–194. doi: 10.1016/j.trf.2019.10.013

Billinghurst, M. (2021). Grand challenges for augmented reality. *Front. Virtual Real.* 2, 12. doi: 10.3389/frvir.2021.578080

Boban, M., Kousaridas, A., Manolakis, K., Eichinger, J., and Xu, W. (2018). Connected roads of the future: use cases, requirements, and design considerations for vehicle-to-everything communications. *IEEE Veh. Technol. Mag.* 13, 110–123. doi: 10.1109/MVT.2017.2777259

Bos, J. E., de Vries, S. C., van Emmerik, M. L., and Groen, E. L. (2010). The effect of internal and external fields of view on visually induced motion sickness. *Appl. Ergon.* 41, 516–521. doi: 10.1016/j.apergo.2009.11.007

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Brooke, J. et al. (1996). Sus-a quick and dirty usability scale. *Usability Eval. Ind.* 189, 4–7.

Cohda Wireless (2017). Australia's first vehicle-to-pedestrian technology trial. (accessed March 17, 2022).

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences (Revised Edition)*. New York, NY: Academic Press.

Colley, M., Belz, J. H., and Rukzio, E. (2021). "Investigating the effects of feedback communication of autonomous vehicles," in *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Washington, DC: Leeds), 263–273.

Colley, M., and Rukzio, E. (2020a). "A design space for external communication of autonomous vehicles," in *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (New York, NY), 212–222.

Colley, M., and Rukzio, E. (2020b). "Towards a design space for external communication of autonomous vehicles," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–8.

Colley, M., Walch, M., Gugenheimer, J., Askari, A., and Rukzio, E. (2020a). "Towards inclusive external communication of autonomous vehicles for pedestrians with vision impairments," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY), 1–14.

Colley, M., Walch, M., and Rukzio, E. (2020b). "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* (New York, NY), 1–9.

Coughlan, J. M., and Miele, J. (2017). "Ar4vi: Ar as an accessibility tool for people with visual impairments," in *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)* (New York, NY), 288–292.

Dey, D., Habibovic, A., Löcken, A., Wintersberger, P., Pfleging, B., Riener, A., et al. (2020a). Taming the ehmi jungle: a classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transp. Res. Interdiscipl. Perspect.* 7, 100174. doi: 10.1016/j.trip.2020.100174

Dey, D., Habibovic, A., Pfleging, B., Martens, M., and Terken, J. (2020b). "Color and animation preferences for a light band ehmi in interactions between automated vehicles and pedestrians," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY), 1–13.

Dey, D., Martens, M., Eggen, B., and Terken, J. (2017). "The impact of vehicle appearance and vehicle behavior on pedestrian interaction with autonomous vehicles," in *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct* (New York, NY), 158–162.

Dey, D., van Vastenhoven, A., Cuijpers, R. H., Martens, M., and Pfleging, B. (2021). "Towards scalable ehmis: designing for av-vru communication beyond one pedestrian," in *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (New York, NY), 274–286.

Epke, M. R., Kooijman, L., and De Winter, J. C. (2021). I see your gesture: a vr-based study of bidirectional communication between pedestrians and automated vehicles. *J. Adv. Transp.* 2021, 5573560. doi: 10.1155/2021/5573560

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using g* power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Grubert, J., Langlotz, T., Zollmann, S., and Regenbrecht, H. (2016). Towards pervasive augmented reality: context-awareness in augmented reality. *IEEE Trans. Visual. Comput. Graph.* 23, 1706–1724. doi: 10.1109/TVCG.2016.2543720

Gruenefeld, U., Weiß, S., Löcken, A., Virgilio, I., Kun, A. L., and Boll, S. (2019). "Vroad: gesture-based interaction between pedestrians and automated vehicles in virtual reality," in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*, 399–404.

Hart, S. G., and Staveland, L. E. (1988). "Development of nasa-tlx (task load index): results of empirical and theoretical research," in *Advances in Psychology*, vol. 52. (North-Holland, MI: Elsevier), 139–183.

Hendricks, D. L., Fell, J. C., Freedman, M., et al. (2001). *The Relative Frequency of Unsafe Driving Acts in Serious Traffic Crashes*. Technical Report, United States. National Highway Traffic Safety Administration.

Hesenius, M., Börsting, I., Meyer, O., and Gruhn, V. (2018). "Don't panic! guiding pedestrians in autonomous traffic with augmented reality," in *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, 261–268.

Holländer, K., Krüger, A., and Butz, A. (2020). "Save the smombies: app-assisted street crossing," in *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services* (New York, NY), 1–11.

Hussein, A., Garcia, F., Armingol, J. M., and Olaverri-Monreal, C. (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)* (Rio de Janeiro: IEEE), 2034–2039.

Jian, J.-Y., Bisantz, A. M., and Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *Int. Journal Cogn. Ergon.* 4, 53–71. doi: 10.1207/S15327566IJCE0401_04

Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L., and Nass, C. (2015). Why did my car just do that? explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *Int. J. Interact. Design Manuf. (IJIDeM)* 9, 269–275. doi: 10.1007/s12008-014-0227-2

Lee, C. K., Yun, I., Choi, J.-H., Ko, S.-j., and Kim, J.-Y. (2013). Evaluation of semi-actuated signals and pedestrian push buttons using a microscopic traffic simulation model. *KSCE J. Civil Eng.* 17, 1749–1760. doi: 10.1007/s12205-013-0040-7

Liu, Z., Pu, L., Meng, Z., Yang, X., Zhu, K., and Zhang, L. (2015). "Pofs: a novel pedestrian-oriented forewarning system for vulnerable pedestrian safety," in *2015 International Conference on Connected Vehicles and Expo (ICCVE)* (Shenzhen: IEEE), 100–105.

Löcken, A., Golling, C., and Riener, A. (2019). "How should automated vehicles interact with pedestrians? a comparative analysis of interaction concepts in virtual reality," in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (New York, NY), 262–274.

Mahadevan, K., Sanoubari, E., Somanath, S., Young, J. E., and Sharlin, E. (2019). "Av-pedestrian interaction design using a pedestrian mixed traffic simulator," in *Proceedings of the 2019 on Designing Interactive Systems Conference* (New York, NY), 475–486.

Mahadevan, K., Somanath, S., and Sharlin, E. (2018). "Communicating awareness and intent in autonomous vehicle-pedestrian interaction," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY), 1–12.

Malik, J., N. Di Napoli Parr, M., Flathau, J., Tang, H., K. Kearney, J., M. Plumert, J., and Rector, K. (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* (New York, NY), 1–12.

Moore, D., Currano, R., Strack, G. E., and Sirkin, D. (2019). "The case for implicit external human-machine interfaces for autonomous vehicles," in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 295–307.

Nguyen, T. T., Holländer, K., Hoggenmueller, M., Parker, C., and Tomitsch, M. (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* (New York, NY), 284–294.

Omeiza, D., Webb, H., Jirotka, M., and Kunze, L. (2021). Explanations in autonomous driving: a survey. *arXiv preprint* arXiv:2103.05154.

Pratticò, F. G., Lamberti, F., Cannavò, A., Morra, L., and Montuschi, P. (2021). Comparing state-of-the-art and emerging augmented reality interfaces for autonomous vehicle-to-pedestrian communication. *IEEE Trans. Veh. Technol.* 70, 1157–1168. doi: 10.1109/TVT.2021.3054312

Prisco, J. (2018). *Illusion of control: Why the world is full of buttons that don't work.* CNN. Available online at: https://edition.cnn.com/style/article/placebo-buttons-design/index.html (accessed March 17, 2022).

Rasouli, A., and Tsotsos, J. K. (2019). "Autonomous vehicles that interact with pedestrians: a survey of theory and practice," *IEEE Trans. Intell. Transp. Syst.* 21, 900–918. doi: 10.1109/TITS.2019.2901817

Ravishankar, K. V. R. and Nair, P. M. (2018). Pedestrian risk analysis at uncontrolled midblock and unsignalised intersections. *J. Traffic Transp. Eng. (English Edition)* 5, 137–147.

Rekimoto, J. (1996). "Transvision: a hand-held augmented reality system for collaborative design," in *Proceeding of Virtual Systems and Multimedia*, vol. 96, 18–20. doi: 10.1016/j.jtte.2017.06.005

Ridel, D., Rehder, E., Lauer, M., Stiller, C., and Wolf, D. (2018). "A literature review on the prediction of pedestrian behavior in urban scenarios," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (Maui, HI: IEEE), 3105–3112.

Riegler, A., Riener, A., and Holzmann, C. (2021). Augmented reality for future mobility: insights from a literature review and hci workshop. *i-com* 20, 295–318. doi: 10.1515/icom-2021-0029

Risto, M., Emmenegger, C., Vinkhuyzen, E., Cefkin, M., and Hollan, J. (2007). "Human-vehicle interfaces: the power of vehicle movement gestures in human road user coordination," in *Driving Assessment: The Ninth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design (2017)* (Manchester, VT), 186–192.

Robert Jr, L. P. (2019). The future of pedestrian-automated vehicle interactions. *XRDS Crossroads ACM Mag. Stud.* 25, 30–33. doi: 10.1145/3313115

SAE (2021). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. *J3016, SAE Int.*

Schneider, S., and Bengler, K. (2020). Virtually the same? analysing pedestrian behaviour by means of virtual reality. *Transp. Res. Part F Traffic Psychol. Behav.* 68, 231–256. doi: 10.1016/j.trf.2019.11.005

Starner, T., Mann, S., Rhodes, B., Levine, J., Healey, J., Kirsch, D., et al. (1997). Augmented reality through wearable computing. *Presence Teleoper. Virt. Environ.* 6, 386–398. doi: 10.1162/pres.1997.6.4.386

Sweeney, M., Pilarski, T., Ross, W. P., and Liu, C. (2018). *Light Output System for a Self-Driving Vehicle.* U.S. Patent 10,160,378.

Tabone, W., De Winter, J., Ackermann, C., Bärgman, J., Baumann, M., Deb, S., et al. (2021a). Vulnerable road users and the coming wave of automated vehicles: expert perspectives. *Transp. Res. Interdiscipl. Perspect.* 9, 100293. doi: 10.1016/j.trip.2020.100293

Tabone, W., Lee, Y. M., Merat, N., Happee, R., and De Winter, J. (2021b). "Towards future pedestrian-vehicle interactions: introducing theoretically-supported ar prototypes," in *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (New York, NY), 209–218.

Tong, Y., and Jia, B. (2019). "An augmented-reality-based warning interface for pedestrians: User interface design and evaluation," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 63 (Los Angeles, CA: SAGE Publications Sage CA), 1834–1838.

Tong, Y., Jia, B., and Bao, S. (2021). An augmented warning system for pedestrians: user interface design and algorithm development. *Applied Sci.* 11, 7197. doi: 10.3390/app11167197

Treat, J. R., Tumbas, N., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R., et al. (1979). *Tri-Level Study of the Causes of Traffic Accidents: Final Report. Executive Summary.* Technical Report, Indiana University, Bloomington, Institute for Research in Public Safety.

Umbrellium (2017). *Make Roads Safer, More Responsive & Dynamic.* Available online at: https://umbrellium.co.uk/case-studies/south-london-starling-cv/ (accessed March 17, 2022).

Urmson, C. P., Mahon, I. J., Dolgov, D. A., and Zhu, J. (2015). *Pedestrian Notifications.* U.S. Patent 8,954,252.

Van Houten, R., Ellis, R., Sanda, J., and Kim, J.-L. (2006). Pedestrian push-button confirmation increases call button usage and compliance. *Transp. Res. Rec.* 1982, 99–103. doi: 10.1177/0361198106198200113

Verstegen, R., Dey, D., and Pfleging, B. (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* (New York, NY), 132–136.

Weber, F., Chadowitz, R., Schmidt, K., Messerschmidt, J., and Fuest, T. (2019). "Crossing the street across the globe: a study on the effects of ehmi on pedestrians in the us, germany and china," in *International Conference on Human-Computer Interaction* (Orlando, FL: Springer), 515–530.

Wiegand, G., Mai, C., Holländer, K., and Hussmann, H. (2019). "Incarar: a design space towards 3d augmented reality applications in vehicles," in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (New York, NY), 1–13.

Wu, X., Miucic, R., Yang, S., Al-Stouhi, S., Misener, J., Bai, S., and Chan, W.-h. (2014). "Cars talk to phones: a dsrc based vehicle-pedestrian safety system," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)* (Vancouver, BC: IEEE), 1–7.

Zileli, S., Davis, S. B., Wu, J., and Zileli, S. (2019). Towards transparency between the autonomous vehicle and the pedestrian. In *DeSForM19 Proceedings.* doi: 10.21428/5395bc37.bd4815e4