The effect of Augmented Reality on Pedestrians' Gaze Patterns and Crossing Probability while Interacting with Automated Vehicles.

Authors

Yue Yang[1]*, Yee Mun Lee[1], Wilbert Tabone[2], Jorge García de Pedro[1], Riender Happee[2], Joost de Winter[2], Natasha Merat[1]

Affiliations

[1] Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, United Kingdom

[2] Department of Cognitive Robotics, Delft University of Technology, 2628 CD Delft, The Netherlands

Corresponding author

*Yue Yang, Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, United Kingdom. Email: y.yang2@leeds.ac.uk

Email addresses

Yue Yang — y.yang2@leeds.ac.uk

Yee Mun Lee — y.m.lee@leeds.ac.uk

Wilbert Tabone — w.tabone@tudelft.nl

Jorge García de Pedro — j.garcia1@leeds.ac.uk

Riender Happee — r.happee@tudelft.nl

Joost de Winter — j.c.f.dewinter@tudelft.nl

Natasha Merat — n.merat@its.leeds.ac.uk

Abstract

The study of Augmented Reality (AR) in transportation has been growing rapidly, and it could be used in communicating the intention of an approaching automated vehicle (AV) to pedestrians. However, it remains unclear whether the adoption of AR increases pedestrians' visual load. This study examined pedestrians' gaze behaviour and crossing decisions when exposed to AR interfaces positioned as a heads-up display (HUD), at the crossing path, or along the AV's travel path. Thirty participants completed trials in a CAVE-based virtual reality (VR) pedestrian lab. We analysed gaze fixations on the vehicle and AR interfaces during the period leading up to crossing initiations. Results showed that, compared to baseline conditions without AR, AR conditions were associated with reduced visual load, indicating that AR did not overburden attention. Interfaces rated as more intuitive and repeated exposures enhanced this effect, though these patterns may also indicate overreliance. Among the different placements, a HUD yielded the greatest decrease in visual load, followed by AR on the crossing path, and then AR along the vehicle's path. Gaze heat maps showed that pedestrians increasingly focused their attention on the vehicle as it approached, regardless of AR locations. Crossing probabilities revealed that in baseline conditions, pedestrians were most likely to cross when the AV was closest and stopped, whereas with AR present, crossings were more likely at greater distances, reflecting earlier recognition of intent. Overall, these findings suggest that AR, if intuitively designed, does not visually overload pedestrians and can support safer crossing decisions, although the potential for overreliance requires further study.

Keywords: Augmented Reality; Automated Vehicles; Pedestrian Safety; Gaze Behaviour; Visual Load; Crossing Decisions

1 Introduction

The introduction of automated vehicles (AVs) leads to a significant transition in transportation, promising many benefits, including a major reduction in accidents involving vulnerable road users by eliminating human errors (Anderson et al., 2016). However, higher-level AVs, which operate without human drivers, are currently unable to effectively communicate their own intentions to surrounding traffic. This limitation can lead to frustrating standoffs, particularly in ambiguous situations where both the AV and other road users are trying to occupy the same space but are uncertain about who has the right of way, such as at unsignalized crossings (Brown et al., 2023; Loke, 2019; Rasouli et al., 2018; Vinkhuyzen & Cefkin, 2016). The absence of a human driver or traffic signals at these crossings prevents clear communication, further complicating the determination of priority and increasing the likelihood of hesitation or hazardous interactions.

External Human-Machine Interfaces (eHMIs) have been proposed as a solution for bridging this communication gap by externally displaying information about AV intentions to pedestrians (Faas et al., 2020; Guo et al., 2022; Hochman et al., 2020; Holländer et al., 2019; Lee et al., 2022; Lyu et al., 2024; Wilbrink et al., 2021). Although eHMIs can help pedestrians make quicker decisions and increase their perceived safety (Faas et al., 2020; Holländer et al., 2019), they face challenges in scalability, particularly for managing multiple interactions simultaneously and effectively communicating across various distances and directions (Colley et al., 2020; Dey et al., 2021; Holländer et al., 2022; Lyu et al., 2024; Wilbrink et al., 2021). These challenges raise concerns about how an AV communicates with specific pedestrians among many road users and the visibility of eHMIs in complex, real-world traffic scenarios (Dey, Habibovic, et al., 2020).

Given these challenges, personalized interaction strategies like Augmented Reality (AR) are being explored as a complementary approach in assisting with communication for pedestrian-AV interactions (Calvi et al., 2020; Matviienko et al., 2022; Tabone et al., 2020, 2021, 2023; Tran et al., 2023). AR allows for simultaneous communication with multiple road users, providing precise, customized visual information to pedestrians (Dey, Habibovic, et al., 2020). By overlaying digital content onto the physical world, this approach offers several benefits, such as resolving language barriers through person-specific feedback

(Tabone et al., 2020), and maintaining users' situational awareness (Tong et al., 2021). Although the use of AR for road user communication may seem futuristic and raise concerns about reliance on costly headsets (Tabone et al., 2020), advancements in wearable AR technology (e.g., Microsoft HoloLens, Google Glass, Apple Vision Pro) are making its adoption in AV-pedestrian communication increasingly feasible.

Despite these potential benefits, there are concerns that AR might overly burden pedestrians with additional visual elements (Tabone et al., 2020). Research in learning and skill acquisition domains has shown that while mobile AR can decrease cognitive load by providing direct information, it can also overwhelm users when presenting excessive information simultaneously (see reviews from Buchner et al., 2022; Suzuki et al., 2024). In road user interactions, pedestrians may experience cognitive and information overload with too many visual cues, posing safety risks (Mahadevan et al., 2018; Moore et al., 2019). Eye-tracking offers a method to measure pedestrians' visual attention, helping to assess whether they are visually overloaded by these cues. Additionally, research examining gaze fixations, defined as periods when the eyes remain relatively still and focus on a specific element, helps gain deeper insights into how pedestrians engage with visual information (Salvucci & Goldberg, 2000). Longer fixation durations may indicate increased visual effort (He & McCarley, 2010; Herten et al., 2017; Jacob & Karn, 2003) or difficulty in processing the visual information (Kotval & Goldberg, 1998; Milton et al., 1950), while shorter fixations suggest quicker information absorption. However, investigations assessing pedestrians' gaze behaviour when exposed to AR interfaces signalling the intentions of Avs have been overlooked.

Research into pedestrians' gaze behaviour can guide the placement and design of AR interfaces (de Winter et al., 2021; Dey et al., 2019), although most current eye-tracking research in AV-pedestrian interactions has been focused on eHMIs (Eisma et al., 2020; Guo et al., 2022; Hochman et al., 2020; Lyu et al., 2024). For instance, Eisma et al. (2020) found that windscreen-mounted eHMIs effectively focused pedestrian gaze, while road projections dispersed gaze patterns and increased visual effort, making them less ideal. Also, this study used a desktop-based 2D simulation setup, which may not have accurately reflected gaze behaviour in a 3D environment. Using a Wizard-of-Oz study, Dey et al. (2019) observed that pedestrians' gaze shifted from the surrounding environment to the car's bumper and gradually to the windshield as the vehicle approached. They recommended distance-dependent eHMIs considering this visual attention pattern from pedestrians. However, Dey et al.'s (2019) study involved the use of stationary pedestrians pressing a button to indicate their crossing intention, rather than making real crossing decisions, possibly limiting insights into natural behaviour in dynamic environments (Te Velde et al., 2005). Additionally, the initial head orientation of the pedestrian, which is known to influence gaze patterns (Tabone et al., 2024), was not controlled. While the above studies suggest that vehicle distance and display placement affect pedestrian gaze, it remains unclear whether AR displays are likely to influence gaze patterns in a similar manner, and whether the pattern is likely to be the same in more dynamic, 3D contexts, when participants' initial attention orientation is more systematically controlled. Addressing these gaps could significantly inform AR placement strategies and potential use cases, as AR can be more versatile in its location compared to eHMIs, which are typically fixed to the vehicle.

In AV-pedestrian interactions, longer gaze durations on AVs are linked to uncertainty about the AV's intentions and increased feelings of danger (Liu et al., 2023). Similarly, longer gaze duration on eHMI designs indicates lower perceived clarity in communicating AV intent to pedestrians (Guo et al., 2022). Research suggests that intuitive eHMI designs can reduce confusion and ease pedestrians' information load (Moore et al., 2019), with repeated exposures fostering greater trust, faster crossing decisions, fewer gaze fixations, and reduced attentional behaviours like head-turning (Faas et al., 2020; Hochman et al., 2020; Yang et al., 2024). Intuitive AR designs may offer similar benefits, potentially streamlining decision-making by enabling pedestrians to assess crossing safety more quickly (Tabone et al., 2024), especially with repeated exposures. This increased efficiency in comprehension could lead to shorter gaze fixation durations on both AV and AR elements in AR-present versus no-AR trials, indicating reduced visual demands. However, the correlation between intuitive design and gaze fixation patterns, particularly with repeated exposures, remains underexplored. Investigating this relationship could significantly inform AR design for safer and more efficient AV-pedestrian interactions.

Additionally, if pedestrians' gaze patterns could be influenced by different AR placements at different AV distances, one can assume that their crossing decisions could also change correspondingly, as gaze behaviour often correlates with decision-making in value-based choice experiments (Anderson, 2013; Gluth et al., 2018, 2020; Krajbich et al., 2010; Krajbich & Rangel, 2011; Shimojo et al., 2003; Thomas et al., 2019). Research has shown that pedestrians presented higher crossing probabilities with the presence of an eHMI communicating the AV's intentions at greater AV distances before fully stopping (Dey, Matviienko, et al., 2020; Lee et al., 2022; Pekkanen et al., 2022; Schneemann & Gohl, 2016). AR could have a similar effect, potentially leading pedestrians to decide to cross earlier, while the AV is still at a greater distance. However, the effect of AR placement on both gaze and the timing of crossing decisions remains unexplored. Investigating this relationship could provide critical insights into where AR should be positioned to optimise AV-pedestrian communication at various distances.

In response to these considerations, our study posed the following research questions:

1. How do different AR locations influence pedestrians' gaze patterns as an AV approaches?
2. How do the location, intuitiveness and repeated encounters of AR influence pedestrians' fixation duration during the crossing task?
3. How do different AR locations influence pedestrians' crossing probabilities at various AV approach distances?

To address these questions, our road crossing study examined pedestrians' gaze behaviour while exposed to a variety of AR concepts, which were proposed in Tabone et al. (2023) and Tabone et al. (2024), in a CAVE-based pedestrians simulator environment.

2 Method

2.1 Participants

Thirty participants were recruited for this study through the University of Leeds Driving Simulator Database, social media and university mailing lists. Among the participants, 20 were males, nine were females, and one was unspecified (age range 22-53 years, M = 31.50, SD = 7.98). All participants were required to be aged 18 and above, possess proficient English language skills, and be free from significant mobility limitations, epilepsy, claustrophobia, or proneness to disorientation. To compensate for taking part in the study (60-90 minutes), each participant received a £15 Amazon gift voucher. The study received ethical approval from the University of Leeds Research Ethics Committee (Ref: LLTRAN-150).

2.2 Apparatus and the virtual environment

The study was conducted in the Highly Immersive Kinematic Experimental Research (HIKER) simulator, a 9 × 4 m CAVE environment at the University of Leeds (as shown in Figure 1). It comprised eight 4K projectors and 10 Vicon Vero 2.2 IR cameras, managed via Vicon Tracker 3.9. The experimental virtual environment, designed in Unity, replicated a residential one-way street featuring a single lane 3.6 meters wide, which was the same as Lee et al. (2022). Eye-tracking data were captured at a frequency of 50 Hz using the Tobii Pro Glasses 2, operated and calibrated with Tobii Controller Software.

Figure 1. A participant in the HIKER lab waits for the start of a trial. In the coordinate system, the 'Y' axis aligns with the participant's height, the 'Z' axis aligns with the pedestrian's intended path, and the 'X' axis aligns with the AV approaching trajectory. The cyan circle in front is an attention attractor used to control the direction of pedestrians' initial focus. It appears randomly, counterbalanced to the left, front, or right of the pedestrian. Pedestrians were required to look at this area, to trigger the start of each trial.

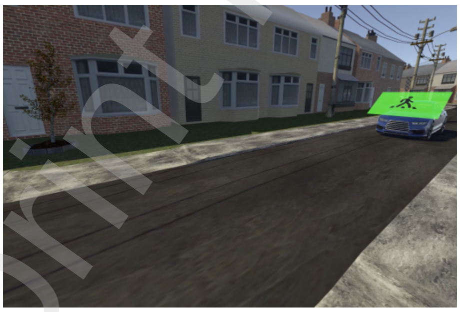
2.3 Study design

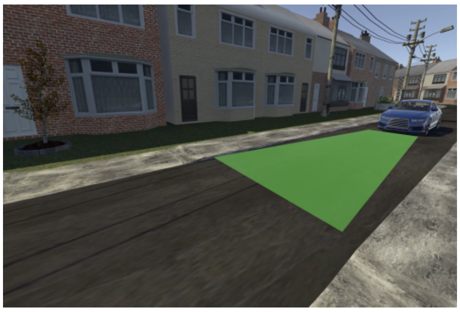
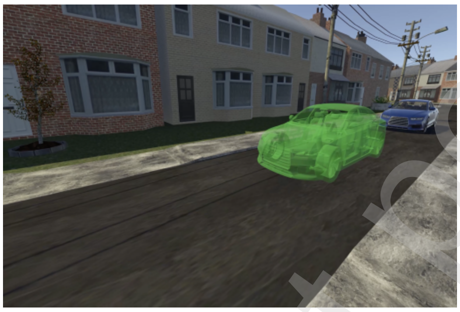
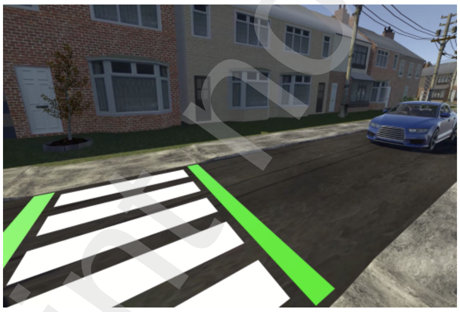
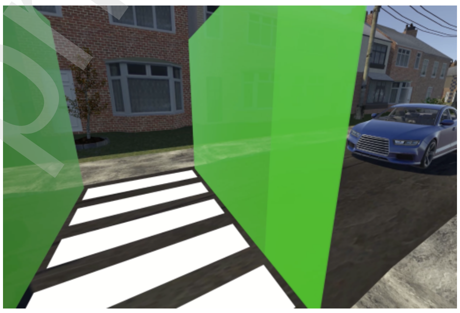
This study builds upon the experiment conducted in Tabone et al. (2024). A within-participant experimental design was implemented, with participants experiencing 10 blocks of 12 trials in each block. There were four independent variables: (i) AR designs (nine AR designs/no AR), (ii) the location of the attention-attractor circle presented before the trial (left/centre/right), (iii) vehicle yielding behaviour (yielding/non-yielding) and (iv) encounter of the yielding trials ($1^{st}/2^{nd}/3^{rd}$).

Each block featured a single AR condition, covering nine AR designs and one baseline without AR. The attention attractor circle was used to simulate real-life situations where pedestrians may be looking in different directions before crossing. Participants were asked to focus on an attention-attractor circle at the start of the trial (the cyan circle shown in Figure 1), located on either the left, centre, or right. They were only allowed to look freely after the circle disappeared. Within each block, participants experienced three trials of yielding AVs and one trial of a non-yielding AV, all approaching from the right. The order of blocks was counterbalanced across participants, and the trials within each block were presented in a randomised order.

The AR designs included in this study are illustrated in Table 1, which was categorised based on their location: (i) Car Path – Four AR designs which were located in the area of the approaching AV following its movement, (ii) Crossing Path – Two AV designs which were located on the crossing path, (iii) Heads-up display (HUD) – Two AR designs which were constantly located in their visual field regardless of head movements. A ninth design in the original study (Tabone et al., 2023, 2024) was excluded from the analysis because it featured a conventional traffic light, which does not fall into either category.

Table 1. Description of AR concepts with categorisations based on their locations

Category	**Design**	**AR concept**
Car Path		**Planes on Vehicle** A plane displayed on the vehicle's windshield area.

		**Conspicuous Looming Planes** A scalable plane that changed size according to the yielding state. It gets smaller in the yielding state.
		**Field of Safe Travel** A projection on the road in front of the vehicle indicating a safe travel area.
		**Phantom Car** A phantom car was displayed to show the vehicle’s predicted future motion.
Crossing Path		**Augmented Zebra Crossing** A zebra crossing was displayed on the crossing path.
		**Virtual Fence** Semi-translucent walls around the zebra crossing with a gate that was opened during the yielding state.

HUD		**Nudge HUD** Text and icons were displayed in the user's field of view.
		**Pedestrian Lights HUD** A traffic light was displayed in the user's field of view.

At the start of each trial, participants stood at Point E (Figure 2.) and fixated on an attention-attractor circle. After one second, the AV departed from Point A at a constant speed of 48 km/h (30 mph). Seven seconds later, it reached Point B (43 m from the participant), triggering the AR interfaces in non-baseline trials.

In yielding trials, the AV began decelerating 0.8 s after Point B (at Point C, 33 m away from the participant), with a rate of 2.99 m/s^2, matching Kaleefathullah et al. (2020). It stopped four seconds later at Point D (3 m from the participant). The attention-attractor circle disappeared precisely 0.2 s after deceleration onset (1 s after Point B), and pedestrians were now allowed to observe the scene and make a crossing decision, as the AV reached 30 meters away.

In non-yielding trials, the attention-attractor circle also disappeared 1 s after Point B, but the AV continued at constant speed.

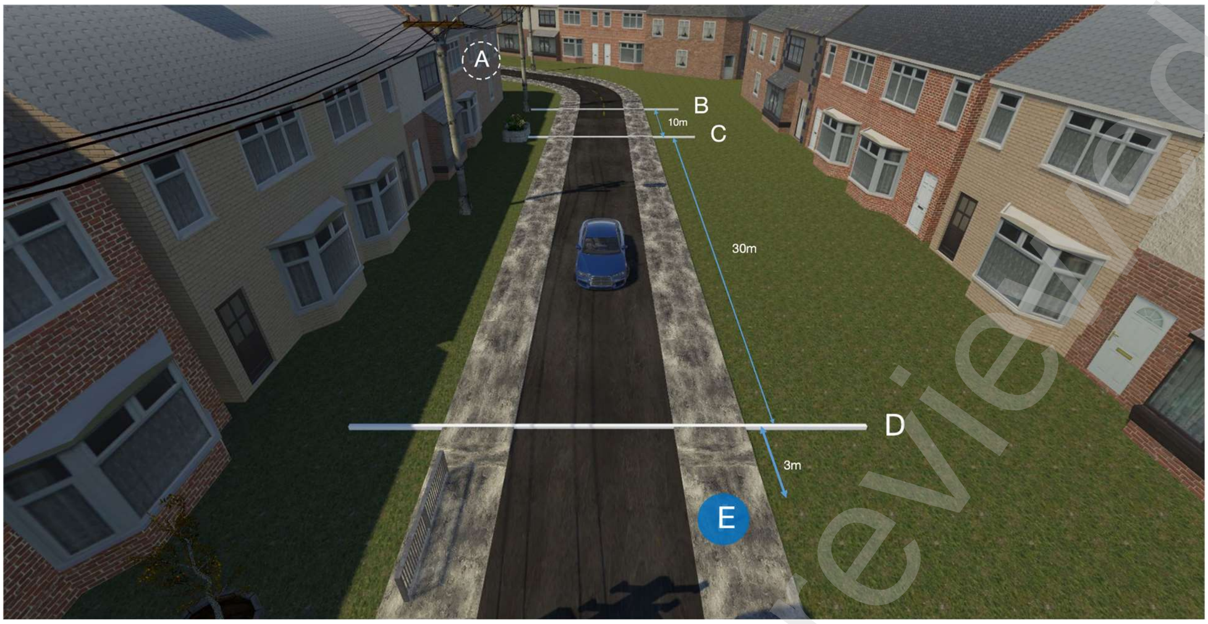


Figure 2. A bird's-eye view of the virtual road layout. Point A marks the starting position of the AV. Point B denotes the activation of the AR interfaces in non-baseline trials. Points C and D represent the onset of deceleration and the stopping point of the AV, respectively, during yielding trials. Point E shows the initial standing position of pedestrians at the beginning of each trial.

2.4 Procedure

Upon arrival at the lab, participants were provided with an information sheet detailing the study and were given a consent form to sign after their queries were addressed. They then completed questionnaires to provide information such as demographics, nationality, and experience with AR/VR, with details reported in Tabone et al. (2024).

Before starting the trials, the eye-tracker was calibrated. Pedestrians were instructed to stand on a blue marker at the beginning of each trial. Once positioned, they initiated the trial by focusing on a stationary, cyan-coloured circle. A continuous one-second gaze on this attention-attracting circle was required to start the trial. If participants' attention deviated, an automatic beeping sound reminded them to refocus on the circle. Successful adherence to this instruction triggered the start of the trial, with the AV entering the simulation from a concealed position. Participants' primary task was to safely cross the virtual road from one curb to another when they felt safe. After providing their answer to the perceived intuitiveness verbally, participants returned to the starting point to begin the next trial.

Two practice trials were conducted before the main experiment: one with a non-yielding vehicle and another with a yielding vehicle. The study began after participants confirmed their understanding of the environment and the task and provided consent to take part.

To measure participants' perceived intuitiveness of the AR, they rated their agreement with the statement: "The interface was intuitive for signalling: 'Please cross the road'" on a scale from 1 (Strongly disagree) to 7 (Strongly agree) after each trial.

Upon completion, participants were thanked for their involvement and received compensation for their time.

2.5 Data analysis

In the current study, non-yielding trials were excluded from further analysis because pedestrians did not initiate crossings in these scenarios, and no learning could be assessed with only a single repetition of non-yielding AVs. As a result, this study analysed 81 trials per participant, covering nine AR conditions (three location-based groups covering eight AR designs plus one baseline), with each condition further subdivided by three initial attention directions and three yielding AVs, totalling 2430 trials. The order of each yielding AV within the initial attention directions and within each AR condition was also labelled as the 1st/2nd/3rd encounter to analyse behaviour changes with repeated exposures.

In this study, the positions of vehicles and pedestrians were consistently logged at a frequency of 120 Hz and pedestrians' gaze data were recorded at 50 Hz. Raw gaze data were selected for analysis from the moment the attention-attractor circle disappeared until either the pedestrian initiated a crossing, or the AV passed, for trials where pedestrians chose not to cross. This period captured the interaction phase between the pedestrian and the AV.

Gaze data were collected using a Tobii Glasses 2 (firmware 1.25.6-citronkola-0; head unit 0.062) mobile eye-tracker, which was operated and calibrated using the Tobii Controller Software v.1.114.20033, with thorough calibration procedures conducted before data collection to ensure accuracy and precision. However, factors such as frequent blinking or missing data could reduce the gaze sample rate. To ensure the quality of gaze data analysis, we identified gaps in the recorded eye-movement data, considering any gap longer than 400 milliseconds as missing data rather than a short interruption like blinking, whi. Trials with more than 30% missing data were excluded, as well as data from Participants 6, 17, and 18, where over 30% of their trials contained more than 30% missing data, resulting in the exclusion of 396 trials (Bindschädel et al., 2022). After further exclusion of 51 trials, where pedestrians did not cross, the final analysis included data from 1983 trials, comprising 1768 AR-present trials and 215 no-AR (Baseline) trials.

2.5.1 Pedestrian gaze patterns

To analyse pedestrians' gaze behaviour during interactions with AVs in a 3D environment, we visualised heat maps of their gaze points on the Y-Z plane (horizontal and vertical visual axes, see coordinate system in Figure 1) as the AV approached at different *Distance Intervals* along the X-axis. This analysis was conducted in a world-referenced coordinate system, assuming gaze positions projected onto a plane perpendicular to the AV's travel direction. Grouping gaze data into intervals, rather than using raw continuous distance, ensures sufficient data points per interval for meaningful visualisation, reducing noise and creating smoother and more interpretable gaze heat maps. This method also highlighted distance-specific shifts in gaze behaviour, making it easier to track attention changes as the AV approached.

Once the attention attractor disappeared, allowing pedestrians to observe the situation and begin their interaction with the AV at a distance of 30 meters, gaze data were grouped into 10-meter *Distance Intervals* for the remaining approach time, with intervals defined as 30–20 m, 20–10 m, and 10–0 m meters away from the pedestrians. These intervals were chosen based on findings from Dey et al. (2019), which suggests significant changes in pedestrians' gaze patterns every 10 meters as a vehicle approaches. Starting the interaction at 30 meters, with a time gap of less than 3 seconds between the pedestrian and the AV, has been shown in previous research to be a situation of higher uncertainty (Tian et al., 2023), necessitating explicit communication mechanisms for right-of-way decisions to ensure safe and smooth interactions.

For each *Distance Interval*, the coordinates of pedestrians' gaze points were visualized on the Y-Z plane, and heat maps were created using Kernel Density Estimation (KDE), a statistical method that smooths data points to produce a continuous density surface. The resulting heat map uses a colour gradient from blue (lower density) to red (higher density) to illustrate how heavily pedestrians scanned the environment, elements of the AV or AR, at different distances as the AV approached. All data processing and visualization were conducted using Python 3.

2.5.2 Change in Fixation Duration (ΔFD)

Longer gaze fixations are associated with higher visual effort and greater difficulty in processing the visual information (He & McCarley, 2010; Herten et al., 2017; Jacob & Karn, 2003; Kotval & Goldberg, 1998; Milton et al., 1950). To investigate how AR would influence pedestrians' visual load, we analysed their gaze fixations on specific areas of interest (AOIs) by tracking the gaze location frame by frame, starting from when the attention-attractor circle disappeared until the pedestrian initiated crossing.

The AOIs investigated in this study were: (1) Car body: The AOI for the car body was defined by its moving 3D spatial boundaries, with the car's centre position (XYZ coordinates) and its dimensions (length, width, and height) being updated for each frame. (2) AR interface: The AOI for the AR interface was represented by a moving plane in the 3D environment, with its centre position and size defined in the virtual space each frame. Gaze points that did not fall within either of these two AOIs were classified as falling into the "other" AOI.

Following the instructions from the Tobii White Paper (Olsen, 2012), raw gaze data were first linearly interpolated for gaps shorter than 75 milliseconds to handle data quality issues and then filtered using a 3-sample moving median filter to smooth high-frequency noise. Fixations were subsequently detected using the I-VT (Identification by Velocity Threshold) algorithm, with a velocity threshold of 100°/s and a minimum fixation duration of 100 milliseconds. Although typical fixation durations can range from 50 to 500 milliseconds, depending on the task (Negi & Mitra, 2020; Rayner, 2009). We adopted the 100 milliseconds threshold in line with established standards (Salvucci & Goldberg, 2000). Adjacent fixations within 0.5° and separated by gaps shorter than 75 milliseconds were merged into a single fixation to account for brief interruptions. The total fixation duration for both the car and AR AOIs was calculated during AR-present trials, and solely on the car during no-AR baseline trials for further analysis.

To assess the impact of AR on pedestrian visual load, we introduced the "Change in Fixation Duration (ΔFD)" metric. We first established each participant's baseline by averaging their total fixation duration on the vehicle in no AR trials, representing visual load without AR. In each AR present trial, we then establish the total fixation duration on both the AR interface and the vehicle. Summing the two AOIs overcame the challenge where, in some AR conditions (e.g., HUDs, Virtual Fence, Car Path), the AR overlapped with the vehicle, making gaze indistinguishable, while in others (e.g., Crossing Path), it competed for attention without overlapping. ΔFD was calculated by subtracting each corresponding participant's baseline fixation time from the total fixation duration on both the AR interface and the vehicle. This metric ensures that ΔFD consistently reflects and quantifies additional attention required by the AR, accounting for individual differences in visual load, and addresses the challenge of distinguishing gaze focus between AR interfaces and the vehicle. A positive ΔFD indicated an increased visual load, while a negative ΔFD suggested a reduced visual load, compared to baseline, during crossing decisions.

To answer the second research question, we conducted a Generalised Linear Mixed Model (GLMM) considering repeated measures analysis (Stroup, 2012) on ΔFD. The model applied a linear distribution with an identity link function and included the following variables: (1) *AR Location* (Car Path, Crossing Path, or HUD), (2) *Intuitiveness Rating* (post-trial scores verbally provided by pedestrians), and (3) *Encounter* (number of interactions within each condition: $1^{st}/2^{nd}/3^{rd}$).

2.5.3 Crossing probabilities

A GLMM was conducted to analyse the likelihood of pedestrians deciding to cross when the AV was at different *Distance Interval* (30–20 m, 20–10 m, and 10–0 m), due to the time sequential nature of these distance intervals (Stroup, 2012). The analysis involved a binary logistic regression with a logit link function, including the main effect of *Distance Interval* and its interaction with *AR Location* (Baseline/ Car Path/ HUD/ Crossing Path) to investigate how pedestrians' crossing probabilities at various AV approach distances are influenced by different AR placements

In this paper, all GLMM analyses included participant as a random effect to account for individual differences, with Bonferroni-adjusted pairwise comparisons for post-hoc analyses. The analysis was conducted using SPSS 28, with a significance level set at $p < .05$.

3 Results

3.1 Gaze heat map

In the Baseline condition without AR, Figure 3 from the left to right shows pedestrians' gaze heat map as the AV approached. When the AV was 30-20 meters away, pedestrians' gaze was more on the environment in front of them (the blob in the left top in the first figure) or on the ground. When the vehicle was closer to 20-10 meters, pedestrians increasingly focused on the car itself. Finally, when the AV was within 10 meters, their gaze concentrated predominantly on the AV, particularly on the windscreen. This gaze pattern, where pedestrians' attention shifted from the environment to the car and driver's seat as the AV approached, was also observed with different AR placements (Figure 4,Figure 5,Figure 6), with slight variations depending on the design.

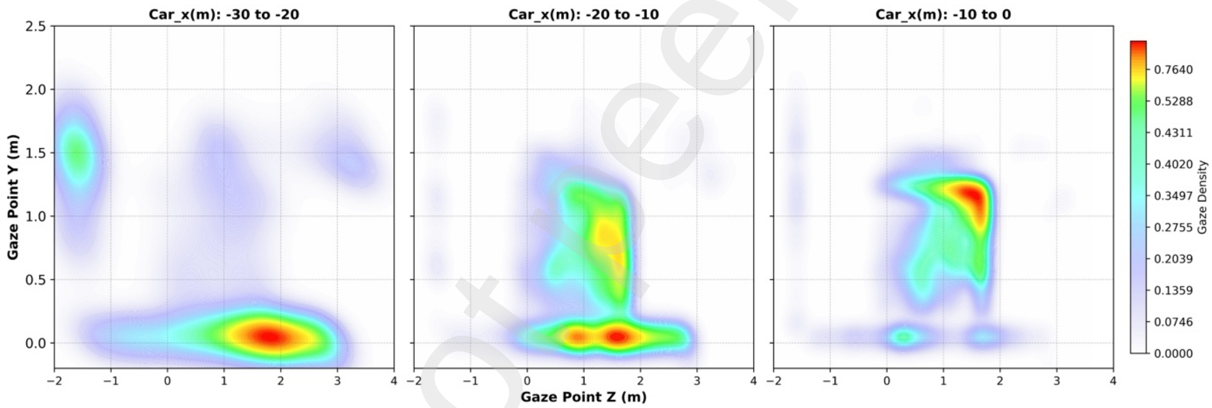


Figure 3. In Baseline trials with no AR concepts, from left to right are pedestrians' gaze heat map on Y-Z plane when the AV's distance to pedestrians (Car_x) was -30 to -20, -20 to -10, and -10 to 0, metres, smoothing using KDE.

With AR on the Car Path (Figure 4a-d), pedestrians' gaze patterns generally resembled the Baseline (Figure 3) when the AV was 30-20 meters, focusing mainly on the environment. However, when a Phantom Car (an AR-generated duplicate of the vehicle indicating its predicted future motion) appeared (Figure 4d), their gaze shifted more towards the vehicle's position in the Y-Z plane (likely focusing on the approaching Phantom Car) between 30 and 20 meters, before concentrating on the windscreen as the AV approached within 20 meters.

In contrast, the other ARs on the Car Path (Figure 4a-c) notably altered gaze behaviour as the AV moved closer, especially between 20-10 meters. Compared to the Baseline (Figure 3, pedestrians focused more on the car and windscreen when the AR was projected onto the windscreen, such as Planes on Vehicle (Figure 4a) and Conspicuous Looming Planes (Figure 4b), with less attention paid to the grill area as the AV was nearly 10 meters away. When AR was projected onto the road, as with the Field of Safe Travel (Figure 4c), pedestrians' attention shifted towards the road between 20-10 meters but became more dispersed across the vehicle and the ground as the AV closed within 10 meters.

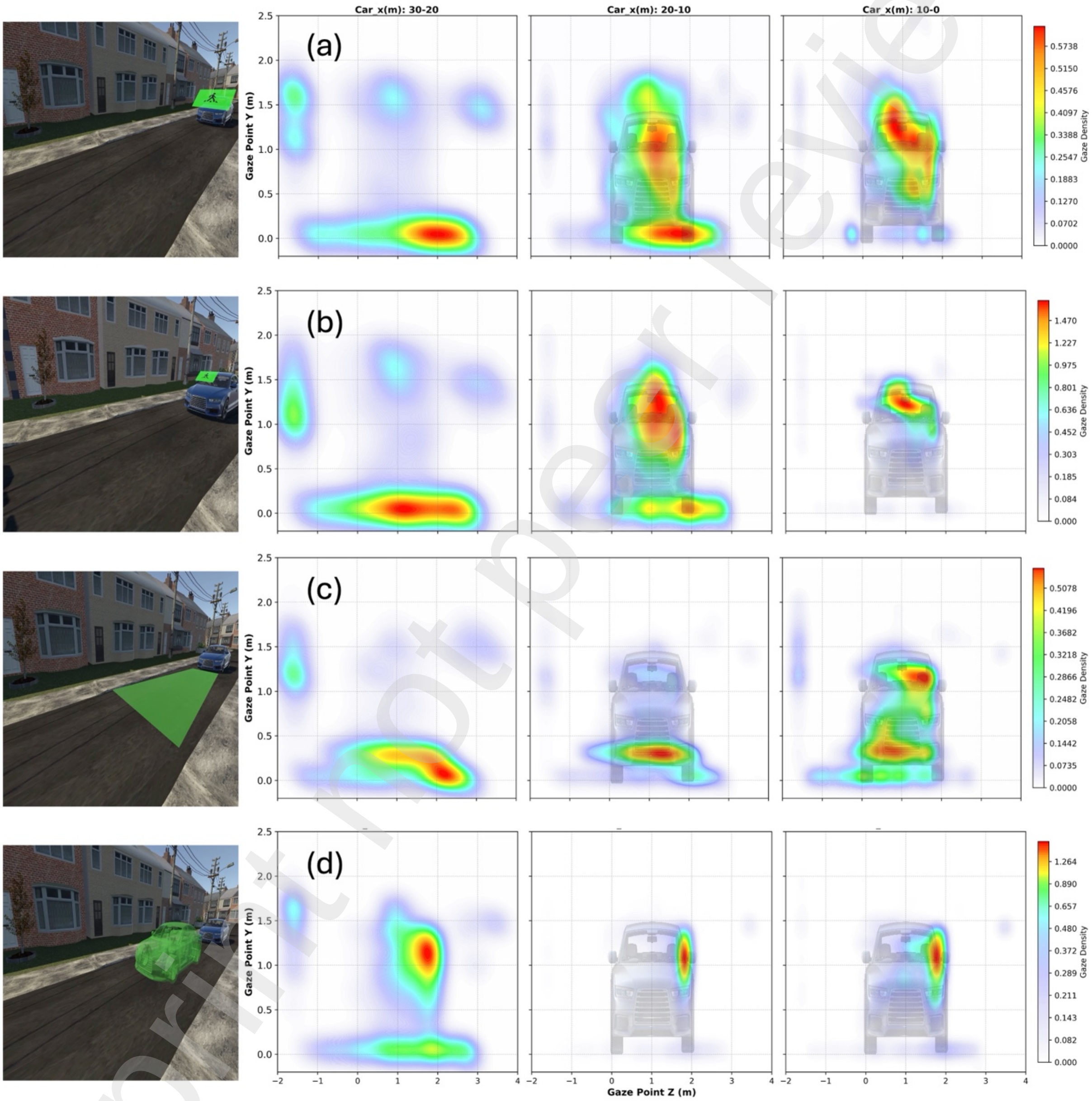


Figure 4. In AR on Car Path, pedestrians' gaze heat maps for designs: (a) Planes on Vehicle, (b) Conspicuous Looming Planes, (c) Field of Safe Travel, and (d) Phantom Car.

In HUD conditions (Figure 5a and b), when the AV was 30-20 and 20-10 meters away, pedestrians focused less on the environment than in Baseline trials, concentrating instead on two areas: the HUD AR and another area likely on the car. As the AV came within 10 meters, their gaze on the windscreen became more dispersed, but there was less focus on the grill compared to the Baseline (Figure 3).

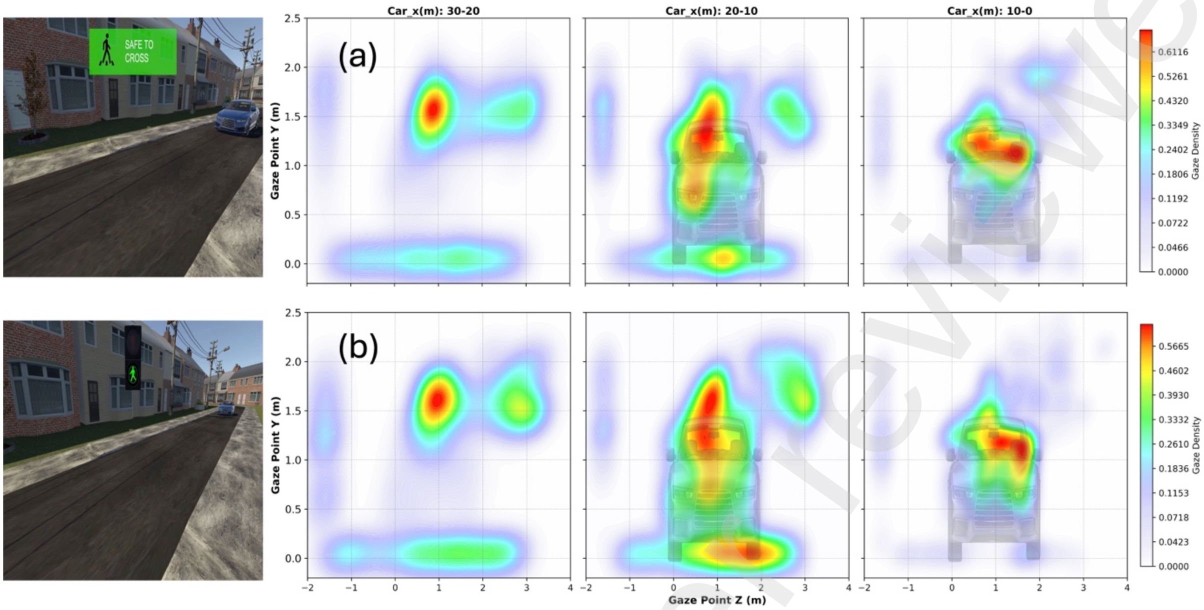


Figure 5. In AR HUD trials, pedestrians' gaze heat maps for designs: (a) Nudge HUD, and (b) Pedestrian Lights HUD.

Regarding AR on Crossing Path (Figure 6a and b), with an Augmented Zebra Crossing (Figure 6a), pedestrians focused more on the ground and less on the car when the AV was beyond 10 meters (30-20 and 20-10), but their gaze became more dispersed across the vehicle and towards the ground as the AV approached within 10 meters, compared to Baseline (Figure 3). With a Virtual Fence added (Figure 6b), pedestrians' gaze remained concentrated on the fence's edge, regardless of the AV's distance.

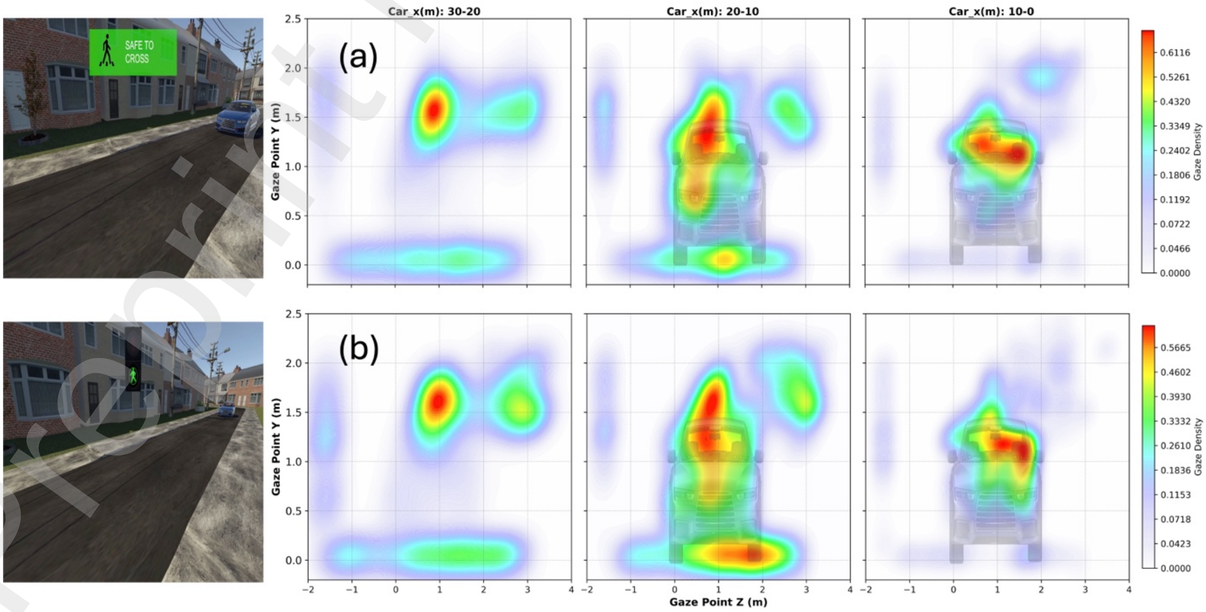


Figure 6. In AR Crossing Path trials, pedestrians' gaze heat maps for designs: (a) Augmented Zebra Crossing, and (b) Virtual Fence.

3.2 Change in Fixation Duration (ΔFD)

A GLMM analysis was conducted to investigate the effects of *AR Location* (Car Path, Crossing Path, or HUD), *Intuitiveness Rating* and *Encounter* on pedestrians' Change in Fixation Duration, with participant included as a random effect. The model revealed significant main effects of *Intuitiveness Rating*, $F (6, 1757) = 21.23$, $p < .001$, *Encounter*, $F (2, 1757) = 8.29$, $p < .001$, and *AR Location*, $F (2, 1757) = 22.74$, $p < .001$.

For *Intuitiveness Rating*, estimated marginal means showed that ΔFD became progressively more negative as ratings increased (from M = 0.13, SE = 0.17 at Rating 1 to M = –0.66, SE = 0.09 at Rating 7), as shown in Figure 7. Sequential Bonferroni-adjusted pairwise comparisons indicated that the most negative ΔFD occurred at Rating 7, which differed significantly from all lower ratings (all ps < .001). Rating 6 (M = –0.61, SE = 0.09) was also significantly more negative than Ratings 5 (p = .018), 4, 3, 2, and 1 (all ps < .001). Rating 5 (M = –0.51, SE = 0.10) was more negative than Ratings 4 (p = .038), 3, 2, and 1 (all ps < .001). Rating 4 (M = –0.40, SE = 0.11) was more negative than Ratings 3 (p = .046), 2, and 1 (both ps < .001). Finally, Rating 3 (M = –0.25, SE = 0.12) was more negative than Rating 1 (p = .004), and Rating 2 (M = –0.09, SE = 0.13) was also more negative than Rating 1 (p = .007). These results suggest that higher perceived intuitiveness of AR interfaces was associated with a greater reduction in visual load compared to baseline.

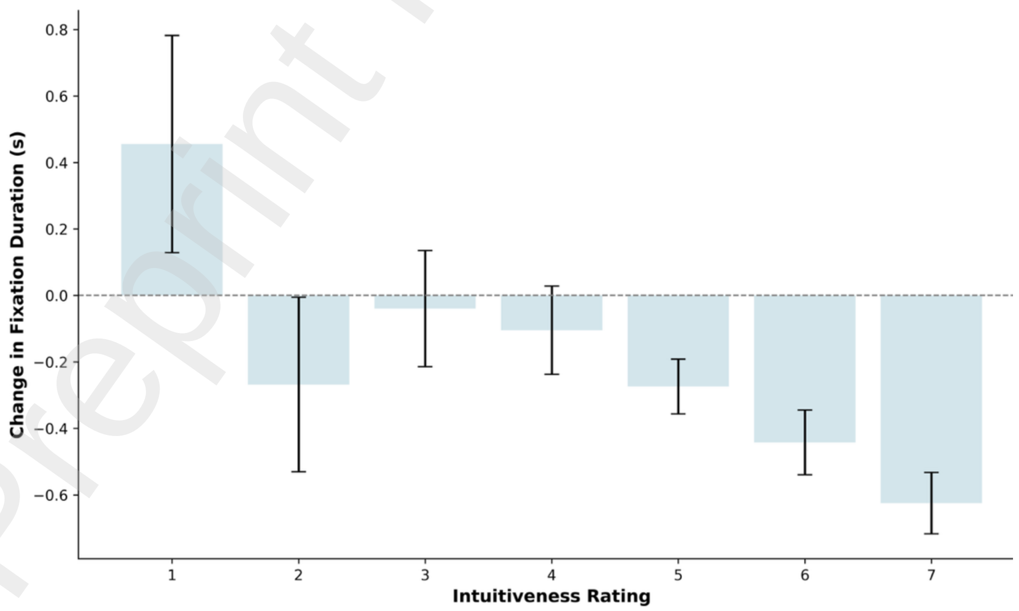


Figure 7. The bar plots and error bars (SE, standard errors, between-subjects) for the impact of Intuitiveness Rating of AR on the Change in Fixation Duration Time.

Car Path showed the least negative ΔFD (M = –0.14, SE = 0.09), significantly less than both HUD (M = –0.33, SE = 0.09; p < .001) and Crossing Path (M = –0.32, SE = 0.09; p < .001), which did not differ from each other (p = .804). Thus, HUD and Crossing Path reduced visual load more than the Car Path interfaces.

For *Encounter*, ΔFD was least negative at the 1st encounter (*M* = –0.18, *SE* = 0.09) and became significantly more negative by the 2nd (*M* = –0.32, *SE* = 0.09; *p* < .001) and 3rd encounters (*M* = –0.30, *SE* = 0.09; *p* = .002). The 2nd and 3rd encounters did not significantly differ (*p* = .515). This indicates that repeated exposure reduced visual load during later encounters with AR.

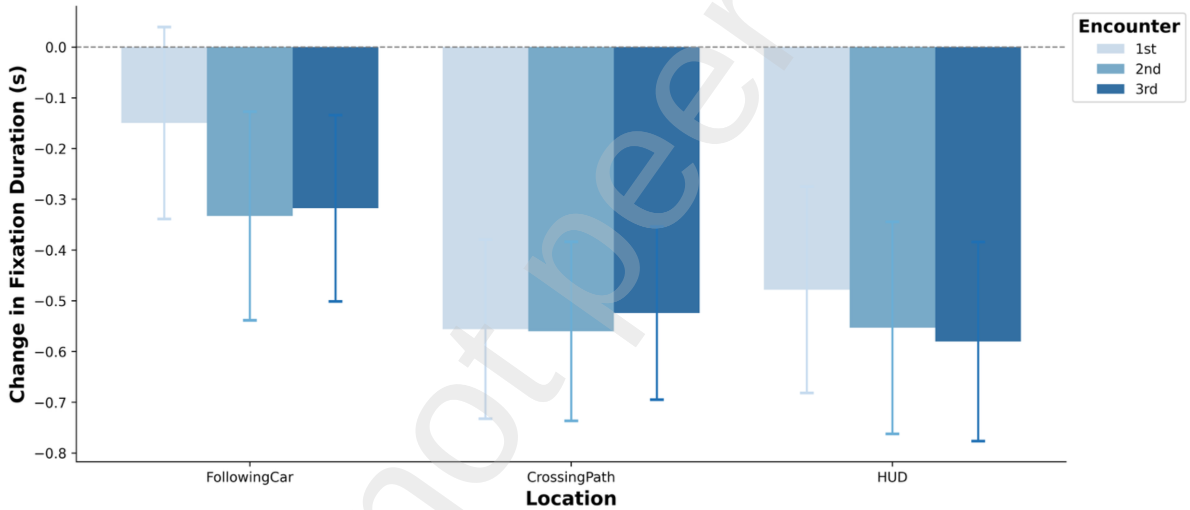


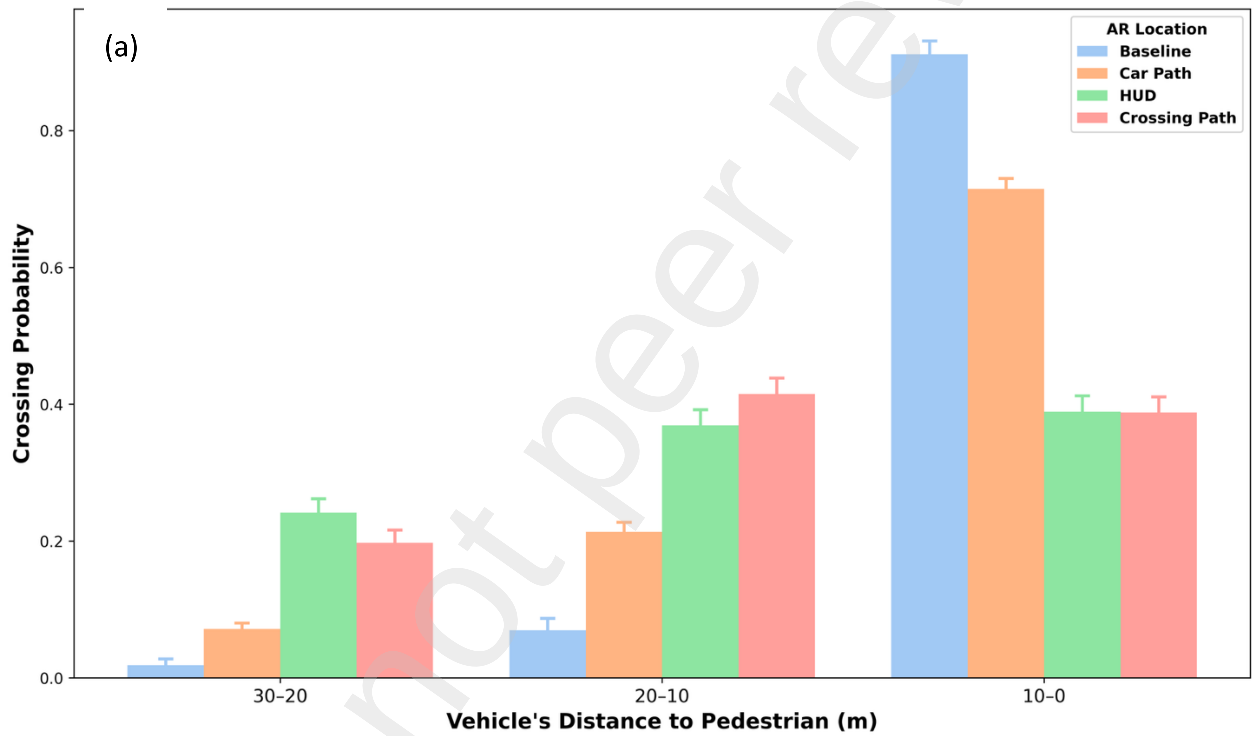
Figure 8. The bar plots and error bars (SE, standard errors, between-subjects) for the impact of AR Location, and the number of Encounter, on the Change in Fixation Duration Time.

3.3 Crossing probability at different AV approach distances

A GLMM was conducted to analyse the likelihood of pedestrians deciding to cross when the AV was at different *Distance Interval* and its interaction with *AR Location*.

The analysis revealed significant effects of *Distance Interval* of AV (F (2,5937) = 237.630, p<.001) and interaction with *AR Location* (F (9, 5937) = 51.553, p < .001) on the probability of crossing.

As shown in **Error! Reference source not found.**a, compared when AV was 30-20 meters away (M = 0.094, SE = 0.012), the likelihood of pedestrians crossing significantly increased as the vehicle approached closer to 20-10 meters (M = 0.233, SE = 0.014, $p < .001$) and 10-0 meters (M = 0.643, SE = 0.016, $p < .001$). Post hoc analysis using LSD showed that crossing probabilities significantly increased as the AV approached closer, from 20-10 meters to 10-0 meters ($p < .001$). However, further post hoc analysis of the interaction effect revealed that this increasing tendency was significant only in the Baseline and AR Car Path conditions. In contrast, crossing probabilities did not differ significantly between the *Distance Interval* of 20-10 meters and 10-0 meters in both the AR Crossing Path and HUD conditions (both $p > .05$).



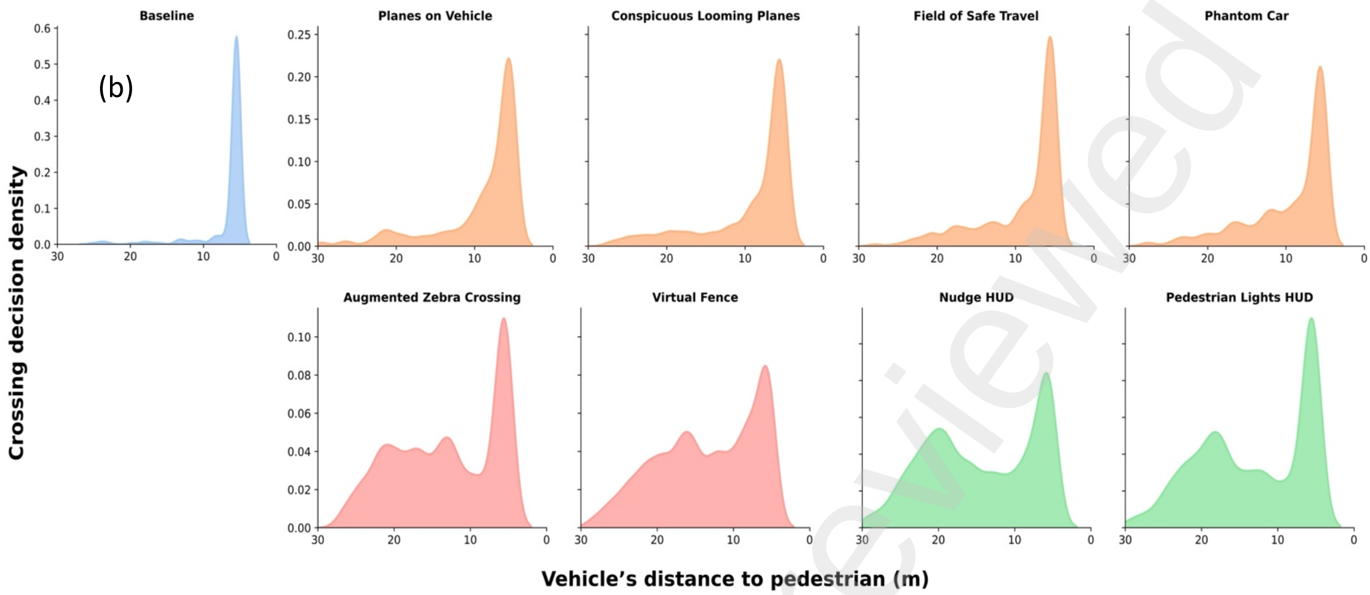


Figure 9.(a) Results from GLMM showing bar plot of pedestrians' crossing decision probabilities across different Distance Interval of AV, clustered by the AR Location. Error bar stands for the standard error. (b) Density plots of crossing probabilities using KDE, depicting their relationship to AV distance in each AR design. The colour scheme matches that of panel (a): blue for Baseline, orange for AR Car Path, green for AR HUD, and red for AR Crossing Path, regarding the AR Location.

When the AV was at both 30-20 and 20-10 meters, the Baseline condition showed significantly lower crossing probabilities compared to the AR Car Path condition (M = 0.019, SE = 0.009 vs. M = 0.072, SE = 0.009; M = 0.070, SE = 0.017 vs. M = 0.214, SE = 0.014; both $p < .001$). Additionally, the crossing probabilities in both the Baseline and AR Car Path conditions were significantly lower than those in the AR Crossing Path condition (M = 0.197, SE = 0.019; M = 0.415, SE = 0.023; both $p < .001$) and the AR HUD condition (M = 0.242, SE = 0.020; M = 0.369, SE = 0.023; both $p < .001$). The differences between AR Crossing Path and AR HUD were not significant in these two intervals (both $p > .05$).

When the AV was within 10 meters away, the Baseline condition showed significantly higher crossing probabilities (M = 0.912, SE = 0.019) compared to the AR Car Path condition (M = 0.715, SE = 0.015, $p < .001$). Both conditions had significantly higher crossing probabilities than the AR Crossing Path condition (M = 0.388, SE = 0.023, both $p < .001$), and the AR HUD condition (M = 0.389, SE = 0.023, both $p < .001$), with no significant difference between AR Crossing Path and AR HUD ($p > .05$).

4 Discussion

This study used a CAVE-based virtual reality pedestrian simulator to investigate pedestrians' gaze patterns and crossing probabilities across different AV approach distances and AR placements. Additionally, this study examined how pedestrians' visual load can be reduced by AR, considering different placements, levels of intuitiveness, and repeated exposures.

Results showed that AR facilitated pedestrians' understanding of the vehicle's intent, as indicated by increased crossing probabilities before the vehicle fully stopped, regardless of AR placement (Figure 9a and b). A bimodal distribution of crossing decisions was observed in previous studies (Dey, Matviienko, et al., 2020; Lee et al., 2022; Pekkanen et al., 2022; Schneemann & Gohl, 2016), where pedestrians were more likely to cross either when the vehicle completed a full stop or at a very far distance. However, this study used a time gap of less than 3 seconds, a kinematic situation characterised by higher uncertainty and more varied interaction patterns (Tian et al., 2023), and pedestrians are more likely to cross when the AV comes to a full stop, which was also observed in baseline trials. However, in such short time gaps, we also identified a bimodal distribution of crossing decisions with AR presence (Figure 9a and b). Similar to the effectiveness of eHMI (Dey, Matviienko, et al., 2020; Lee et al., 2022; Madigan et al., 2023), AR enabled pedestrians to interpret the vehicle's intent earlier, promoting crossing before the vehicle fully stopped. Among the different AR placements, AR on the Crossing Path and HUD showed more effective than Car Path at greater distances, likely due to their higher visibility when the AV was further, which is known to enhance the effectiveness of eHMIs in influencing crossing decisions at various AV approach distances (Dey, Habibovic, et al., 2020; Lee et al., 2022).

Gaze heat maps (Figure 3, Figure 4, Figure 5, Figure 6) revealed a distinct pattern as the vehicle approached, consistent with eye-tracking studies involving manually driven vehicles (de Winter et al., 2021; Dey et al., 2019). When the vehicle was distant, pedestrians primarily scanned the environment or focused on the crossing path or road surface ahead of the vehicle. This behaviour likely occurred because a distant car poses no immediate threat. However, when an immediate threat appeared, such as a fast-approaching phantom car ahead of the AV, pedestrians' attention was captured immediately (Figure 4d). As the AV drew closer, their gaze shifted noticeably from the road to the main body of the vehicle, and eventually to the windshield.

Interestingly, even in the absence of drivers in AVs, pedestrians continued to focus on the windshield, possibly trying to gather information or establish eye contact (Onkhar et al., 2022). A concentrated gaze density was observed on the left seat, which, given that the driver's seat is on the right in the UK, could be attributed to the looming effect (Tian et al., 2022), making the left seat appear closer to pedestrians to be focused on. This suggests that pedestrians' gaze patterns may differ in situations with a driver present, different driving directions, and left-hand or right-hand drive, as well as the road segments and real-world conditions. These dynamics and complexities warrant further investigation.

Additionally, pedestrians' gaze patterns were significantly altered in AR Crossing Path (Figure 6) and HUD (Figure 5) trials compared to Baseline (Figure 3) when the AV was 30 to 20 meters away. In contrast, significant changes in pedestrians' gaze patterns in Car Path (Figure 4) trials were observed when the AV was within 20 meters. This finding highlights AR's potential advantage in AV-pedestrian communication, offering greater versatility in its location and visibility than eHMI attached to or projected from the AV. AR on Crossing Path and HUD are particularly useful when the AV is further away, while AR on Car Path is more effective when the AV is closer.

In this study, AR HUD facilitated earlier crossing decisions (Figure 9) and led to the greatest decrease in fixation duration compared to baseline (Figure 8), aligning with Tabone et al. (2023) findings, where a HUD display was preferred over cues projected on the road or on the approaching vehicle. However, the HUD seemed to distract pedestrians when the AV was more than 10 meters away, as they looked aside to avoid it (Figure 5). Peereboom et al. (2024) found similar results, where HUD received lower ratings and was less preferred compared to baseline, potentially causing discomfort, especially at close distances. In this study, HUD was most effective in influencing crossing decisions when the AV was 30 to 20 meters away, suggesting that such AR is most appropriate for situations involving distant AVs.

Previous research suggested that embedding AR in the environment could divide pedestrians' attention from the oncoming vehicle to the road instead (Peereboom et al., 2024; Tabone et al., 2023). However, this study showed that this was not the case when the AV was farther away. Even with an Augmented Zebra Crossing on the Crossing Path (Figure 6a), pedestrians' attention pattern did not change much compared to Baseline (Figure 3) when the AV was 30 to 10 meters away, as they were not focused on the vehicle during this phase. On the other hand, Virtual Fence led to a more concentrated gaze patterns from pedestrians (Figure 6b). This suggests that AR on the Crossing Path may be best used when the AV is farther away to avoid distracted attention as it approaches.

Some research has highlighted that visually demanding tasks and distractions pose significant risks to pedestrians (Tapiro et al., 2020), suggesting that adding external interfaces could exacerbate these issues, especially when pedestrians rely primarily on kinematic cues from vehicles to make crossing decisions (de Winter & Dodou, 2022; Li et al., 2018). However, our findings reveal that the presence of AR concepts did not increase the fixation duration and visual demands in AV-pedestrian communication, even with the introduction of additional external messages, provided these are intuitively designed (Figure 7). As illustrated in Figure 7, the ΔFD was negative when the AR concept was perceived as intuitive, indicating a reduction in visual load compared to the baseline scenario with no external messages. However, a poorly designed AR interface can impose additional visual effort, potentially undermining pedestrian safety, whereas intuitive designs reduce processing demands and support safer crossing behaviour. This aligns with recommendations from other eHMI studies advocating for messages that are both intuitive and familiar to pedestrians (de Clercq et al., 2019; Hensch et al., 2019; Lee et al., 2022). This result demonstrates the potential benefits of integrating AR in AV-pedestrian communication and underscores the importance of creating clear and intuitive interfaces for safe and efficient pedestrian engagements. However, this research only involves simple one-to-one interactions, leaving uncertainty about how the presence of multiple vehicles might impact visual load for some designs, particularly those associated with the AV (Car Path). For instance, with multiple AVs, each vehicle could project different information, potentially overwhelming pedestrians with competing signals. In contrast, HUD and Crossing Path designs are intended to provide consistent, situationally aware guidance that doesn't change with each individual vehicle. This discrepancy in AR concepts could have a significant effect on visual load, especially as pedestrians attempt to process information from multiple sources simultaneously.

Moreover, repeated exposure to AR interfaces significantly enhanced their effectiveness, supporting the notion of a positive learning curve in AR adoption across all three placements (Figure 8). As participants became more familiar with the interfaces, their ΔFD decreased, suggesting that regular interaction with AR could boost pedestrian confidence and safety over time, even after just one exposure. This observation is consistent with studies indicating that pedestrians can quickly adapt to novel types of eHMIs after several encounters (de Clercq et al., 2019; Eisele & Petzoldt, 2022; Lee et al., 2022; Yang et al., 2024).

While our findings indicate that intuitive AR can reduce visual load (negative ΔFD) and promote earlier crossing decisions, these benefits also raise concerns about potential over-reliance. Prior work has shown that pedestrians sometimes prioritise external messages over vehicle kinematics, stepping into the road even when signals are misleading or incongruent with motion cues (Holländer et al., 2019; Kaleefathullah et al., 2020). Repeated exposure may further reinforce this dependency, as pedestrians adapt to AR cues and reduce head checks or monitoring of the vehicle (Yang et al., 2024). Such over-trust could undermine safety in real traffic, especially in multi-vehicle contexts where competing or inconsistent AR projections may overwhelm attention. To mitigate these risks, AR should complement rather than substitute kinematic information, with intuitive interfaces designed to support safety crossing decisions. This aspect warrants further investigation to ensure the safe application of AR technologies in pedestrian environments.

5 Limitations and Future Work

While this study offers insights for designing AR interfaces in AV-pedestrian communication, it also has limitations that suggest areas for future research. First of all, while ARs have the advantage of communicating with multiple road users over eHMIs, this research only investigates one pedestrian interaction with one AV at a time. Future research could explore AR's role in more complex interactions.

The experimental context was simplified, focusing on an open, straight road and future research can be built on a complex traffic scenario such as intersections or roundabouts, as well as different road infrastructure such as zebra crossings (Madigan et al., 2023; Yang et al., 2024). Additionally, further research can extend this study under different kinematic situations with different driving behaviours and time gaps, which may identify a different role of explicit communication in varying implicit conditions (Dey, Matviienko, et al., 2020; Lee et al., 2022; Madigan et al., 2023). Furthermore, the homogeneity of participant demographics, such as age and gender, which are known to influence attention allocation (Tapiro et al., 2016), can be further explored to propose more personalised AR. Future research should aim to test these AR interfaces in more varied and dynamic outdoor scenarios to validate their effectiveness across different pedestrian populations and urban settings.

6 Conclusion

This study showcases the promising role of AR in enhancing pedestrian safety and decision-making in AV contexts, emphasizing the importance of intuitive, familiar, and repeatedly exposed AR interfaces in reducing visual load. The study also indicates that AR can be more useful when the AV is farther away and there is more uncertainty about its intents. However, it is still crucial to continue refining these technologies through real-world testing and broader user engagement to ensure that they meet the varied needs of all pedestrians in increasingly automated urban environments.

Acknowledgement

This research was supported by a grant from the European Union's Horizon 2020 research and innovation programme, funded under the Marie Skłodowska-Curie actions, grant agreement number 860410.

References

Anderson, B. A. (2013). A value-driven mechanism of attentional selection. *Journal of Vision*, *13*(3). https://doi.org/10.1167/13.3.7

Anderson, J., Kalra, N., Stanley, K., Sorensen, P., Samaras, C., & Oluwatola, O. (2016). Autonomous Vehicle Technology: A Guide for Policymakers. In *Autonomous Vehicle Technology: A Guide for Policymakers*. https://doi.org/10.7249/rr443-2

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

Brown, B., Broth, M., & Vinkhuyzen, E. (2023, April 19). The Halting problem: Video analysis of self-driving cars in traffic. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3544548.3581045

Buchner, J., Buntins, K., & Kerres, M. (2022). The impact of augmented reality on cognitive load and performance: A systematic review. In *Journal of Computer Assisted Learning* (Vol. 38, Issue 1, pp. 285–303). John Wiley and Sons Inc. https://doi.org/10.1111/jcal.12617

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

Colley, M., Walch, M., & Rukzio, E. (2020, April 25). Unveiling the lack of scalability in research on external communication of autonomous vehicles. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3334480.3382865

de Clercq, K., Dietrich, A., Núñez Velasco, J. P., de Winter, J., & Happee, R. (2019). External Human-Machine Interfaces on Automated Vehicles: Effects on Pedestrian Crossing Decisions. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *61*(8), 1353–1370. https://doi.org/10.1177/0018720819836343

de Winter, J., Bazilinskyy, P., Wesdorp, D., de Vlam, V., Hopmans, B., Visscher, J., & Dodou, D. (2021). How do pedestrians distribute their visual attention when walking through a parking garage? An eye-tracking study. *Ergonomics*, *64*(6), 793–805. https://doi.org/10.1080/00140139.2020.1862310

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

Dey, D., Habibovic, A., Löcken, A., Wintersberger, P., Pfleging, B., Riener, A., Martens, M., & Terken, J. (2020). Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transportation Research Interdisciplinary Perspectives*, *7*. https://doi.org/10.1016/j.trip.2020.100174

Dey, D., Matviienko, A., Berger, M., Pfleging, B., Martens, M., & Terken, J. (2020). Communicating the Intention of an Automated Vehicle to Pedestrians: the Contributions of eHMI and Vehicle Behavior. *Information Technology*, *Special Issue: Automotive User Interfaces in the Age of Automation*. https://doi.org/10.1515/ITIT-2020-0025

Dey, D., Van Vastenhoven, A., Cuijpers, R. H., Martens, M., & Pfleging, B. (2021). Towards scalable eHMIs: Designing for AV-VRU communication beyond one pedestrian. *Proceedings - 13th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2021*, 274–286. https://doi.org/10.1145/3409118.3475129

Dey, D., Walker, F., Martens, M., & Terken, J. (2019). Gaze patterns in pedestrian interaction with vehicles: Towards effective design of external human-machine interfaces for automated vehicles. *Proceedings - 11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019*, 369–378. https://doi.org/10.1145/3342197.3344523

Eisele, D., & Petzoldt, T. (2022). Effects of traffic context on eHMI icon comprehension. *Transportation Research Part F: Traffic Psychology and Behaviour*, *85*, 1–12. https://doi.org/10.1016/j.trf.2021.12.014

Eisma, Y. B., van Bergen, S., ter Brake, S. M., Hensen, M. T. T., Tempelaar, W. J., & de Winter, J. C. F. (2020). External human-machine interfaces: The effect of display location on crossing intentions and eye movements. *Information (Switzerland)*, *11*(1). https://doi.org/10.3390/info11010013

Faas, S. M., Kao, A. C., & Baumann, M. (2020). A Longitudinal Video Study on Communicating Status and Intent for Self-Driving Vehicle A- Pedestrian Interaction. *Conference on Human Factors in Computing Systems - Proceedings*, 1–14. https://doi.org/10.1145/3313831.3376484

Gluth, S., Kern, N., Kortmann, M., & Vitali, C. L. (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour*, *4*(6), 634–645. https://doi.org/10.1038/s41562-020-0822-0

Gluth, S., Spektor, M. S., & Rg Rieskamp, J. (2018). *Value-based attentional capture affects multi-alternative decision making*. https://doi.org/10.7554/eLife.39659.001

Guo, F., Lyu, W., Ren, Z., Li, M., & Liu, Z. (2022). A Video-Based, Eye-Tracking Study to Investigate the Effect of eHMI Modalities and Locations on Pedestrian–Automated Vehicle Interaction. *Sustainability (Switzerland)*, *14*(9). https://doi.org/10.3390/su14095633

He, J., & McCarley, J. S. (2010). Executive working memory load does not compromise perceptual processing during visual search: Evidence from additive factors analysis. *Attention, Perception, and Psychophysics*, *72*(2), 308–316. https://doi.org/10.3758/APP.72.2.308

Hensch, A. C., Neumann, I., Beggiato, M., Halama, J., & Krems, J. F. (2019). Effects of a light-based communication approach as an external HMI for Automated Vehicles - A Wizard-of-Oz Study. *Transactions on Transport Sciences*, *10*(2), 18–32. https://doi.org/10.5507/TOTS.2019.012

Herten, N., Otto, T., & Wolf, O. T. (2017). The role of eye fixation in memory enhancement under stress – An eye tracking study. *Neurobiology of Learning and Memory*, *140*, 134–144. https://doi.org/10.1016/j.nlm.2017.02.016

Hochman, M., Parmet, Y., & Oron-Gilad, T. (2020). Pedestrians' Understanding of a Fully Autonomous Vehicle's Intent to Stop: A Learning Effect Over Time. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.585280

Holländer, K., Hoggenmüller, M., Gruber, R., Völkel, S. T., & Butz, A. (2022). Take It to the Curb: Scalable Communication Between Autonomous Cars and Vulnerable Road Users Through Curbstone Displays. *Frontiers in Computer Science*, *4*. https://doi.org/10.3389/fcomp.2022.844245

Holländer, K., Wintersberger, P., & Butz, A. (2019). Overtrust in external cues of automated vehicles: An experimental investigation. *Proceedings - 11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019*, *September*, 211–221. https://doi.org/10.1145/3342197.3344528

Jacob, R. J. K., & Karn, K. S. (2003). Eye Tracking in Human-Computer Interaction and Usability Research. In *The Mind's Eye* (Vol. 42, Issue 5, pp. 573–605). Elsevier. https://doi.org/10.1016/B978-044451020-4/50031-1

Kaleefathullah, A. A., Merat, N., Lee, Y. M., Eisma, Y. B., Madigan, R., Garcia, J., & Winter, J. de. (2020). External Human–Machine Interfaces Can Be Misleading: An Examination of Trust Development and Misuse in a CAVE-Based Pedestrian Simulation Environment. *Human Factors*, *November 2020*. https://doi.org/10.1177/0018720820970751

Kotval, X. P., & Goldberg, J. H. (1998). Eye Movements and Interface Component Grouping: An Evaluation Method. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *42*(5), 486–490. https://doi.org/10.1177/154193129804200509

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. https://doi.org/10.1038/nn.2635

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(33), 13852–13857. https://doi.org/10.1073/pnas.1101328108

Lee, Y. M., Madigan, R., Uzondu, C., Garcia, J., Romano, R., Markkula, G., & Merat, N. (2022). Learning to interpret novel eHMI: The effect of vehicle kinematics and eHMI familiarity on pedestrian' crossing behavior. *Journal of Safety Research*, *80*(January 2022), 270–280. https://doi.org/10.1016/j.jsr.2021.12.010

Li, Y., Dikmen, M., Hussein, T. G., Wang, Y., & Burns, C. (2018). To cross or not to cross: Urgency-based external warning displays on autonomous vehicles to improve pedestrian crossing safety. *Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018*, 188–197. https://doi.org/10.1145/3239060.3239082

Liu, H., Hirayama, T., Morales Saiki, L. Y., & Murase, H. (2023). Implicit Interaction with an Autonomous Personal Mobility Vehicle: Relations of Pedestrians' Gaze Behavior with Situation Awareness and Perceived Risks. *International Journal of Human-Computer Interaction*, *39*(10), 2016–2032. https://doi.org/10.1080/10447318.2022.2073006

Loke, S. W. (2019). Cooperative Automated Vehicles: A Review of Opportunities and Challenges in Socially Intelligent Vehicles beyond Networking. *IEEE Transactions on Intelligent Vehicles*, *4*(4), 509–518. https://doi.org/10.1109/TIV.2019.2938107

Lyu, W., Zhang, W., Wang, X., Ding, Y., & Yang, X. (2024). Pedestrians' responses to scalable automated vehicles with different external human-machine interfaces: Evidence from a video-based eye-tracking experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *103*, 112–127. https://doi.org/10.1016/j.trf.2024.04.005

Madigan, R., Mun Lee, Y., Lyu, W., Horn, S., Garcia de Pedro, J., & Merat, N. (2023). Pedestrian interactions with automated vehicles: Does the presence of a zebra crossing affect how eHMIs and movement patterns are interpreted? *Transportation Research Part F: Traffic Psychology and Behaviour*, *98*, 170–185. https://doi.org/10.1016/j.trf.2023.09.003

Mahadevan, K., Somanath, S., & Sharlin, E. (2018). Communicating awareness and intent in autonomous vehicle-pedestrian interaction. *Conference on Human Factors in*

Computing Systems - Proceedings, *2018-April*. https://doi.org/10.1145/3173574.3174003

Matviienko, A., Müller, F., Schön, D., Seesemann, P., Günther, S., & Mühlhäuser, M. (2022, April 29). BikeAR: Understanding Cyclists' Crossing Decision-Making at Uncontrolled Intersections using Augmented Reality. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3491102.3517560

Milton, J. L., Jones, R. E., & Fitts, P. M. (1950). *Eye Fixations of Aircraft Pilots: Frequency, duration, and sequence fixations when flying selected maneuvers during instrument and visual flight conditions* (Issue 6018). United States, Air Force, Air Materiel Command.

Moore, D., Currano, R., Strack, G. E., & Sirkin, D. (2019). The case for implicit external human-machine interfaces for autonomous vehicles. *Proceedings - 11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019*, 295–307. https://doi.org/10.1145/3342197.3345320

Negi, S., & Mitra, R. (2020). Fixation duration and the learning process: an eye tracking study with subtitled videos. *Journal of Eye Movement Research*, *13*(6), 1–15. https://doi.org/10.16910/jemr.13.6.1

Olsen, A. (2012). The Tobii I-VT fixation filter. *Tobii Technology*, *21*(4–19), 5.

Onkhar, V., Bazilinskyy, P., Dodou, D., & de Winter, J. C. F. (2022). The effect of drivers' eye contact on pedestrians' perceived safety. *Transportation Research Part F: Traffic Psychology and Behaviour*, *84*, 194–210. https://doi.org/10.1016/j.trf.2021.10.017

Peereboom, J., Tabone, W., Dodou, D., & de Winter, J. (2024). Head-locked, world-locked, or conformal diminished-reality? An examination of different AR solutions for pedestrian safety in occluded scenarios. *Virtual Reality*, *28*(2). https://doi.org/10.1007/s10055-024-01017-9

Pekkanen, J., Giles, O. T., Lee, Y. M., Madigan, R., Daimon, T., Merat, N., & Markkula, G. (2022). Variable-Drift Diffusion Models of Pedestrian Road-Crossing Decisions.

Computational Brain and Behavior, *5*(1), 60–80. https://doi.org/10.1007/s42113-021-00116-z

Rasouli, A., Kotseruba, I., & Tsotsos, J. K. (2018). Towards Social Autonomous Vehicles: Understanding Pedestrian-Driver Interactions. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2018-Novem*, 729–734. https://doi.org/10.1109/ITSC.2018.8569324

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Salvucci, D. D., & Goldberg, J. H. (2000). *Identifying Fixations and Saccades in Eye-Tracking Protocols*.

Schneemann, F., & Gohl, I. (2016). Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments. *IEEE Intelligent Vehicles Symposium, Proceedings*, *2016-Augus*(Iv), 38–43. https://doi.org/10.1109/IVS.2016.7535361

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322. https://doi.org/10.1038/nn1150

Stroup, W. W. (2012). *Generalized linear mixed models: modern concepts, methods and applications*. CRC press.

Suzuki, Y., Wild, F., & Scanlon, E. (2024). Measuring cognitive load in augmented reality with physiological methods: A systematic review. In *Journal of Computer Assisted Learning* (Vol. 40, Issue 2, pp. 375–393). John Wiley and Sons Inc. https://doi.org/10.1111/jcal.12882

Tabone, W., De Winter, J., Deb, S., & Habibovic, A. (2020). *Vulnerable Road Users and The Coming Wave of Automated Vehicles: Expert Perspectives*. *October*. https://www.researchgate.net/publication/344657741

Tabone, W., Happee, R., García, J., Lee, Y. M., Lupetti, M. L., Merat, N., & de Winter, J. (2023). Augmented reality interfaces for pedestrian-vehicle interactions: An online study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *94*, 170–189. https://doi.org/10.1016/j.trf.2023.02.005

Tabone, W., Happee, R., Yang, Y., Sadraei, E., García de Pedro, J., Lee, Y. M., Merat, N., & de Winter, J. (2024). Immersive insights: evaluating augmented reality interfaces for pedestrians in a CAVE-based experiment. *Frontiers in Virtual Reality*, *5*. https://doi.org/10.3389/frvir.2024.1353941

Tabone, W., Lee, Y. M., Merat, N., Happee, R., & De Winter, J. (2021). Towards future pedestrian-vehicle interactions: Introducing theoretically-supported AR prototypes. *Proceedings - 13th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2021*, 209–218. https://doi.org/10.1145/3409118.3475149

Tapiro, H., Borowsky, A., Oron-Gilad, T., & Parmet, Y. (2016). Where do older pedestrians glance before deciding to cross a simulated two-lane road? A pedestrian simulator paradigm. *Proceedings of the Human Factors and Ergonomics Society*, 11–15. https://doi.org/10.1177/1541931213601003

Tapiro, H., Oron-Gilad, T., & Parmet, Y. (2020). Pedestrian distraction: The effects of road environment complexity and age on pedestrian's visual attention and crossing behavior. *Journal of Safety Research*, *72*, 101–109. https://doi.org/10.1016/j.jsr.2019.12.003

Te Velde, A. F., Van Der Kamp, J., Barela, J. A., & Savelsbergh, G. J. P. (2005). Visual timing and adaptive behavior in a road-crossing simulation study. *Accident Analysis and Prevention*, *37*(3), 399–406. https://doi.org/10.1016/j.aap.2004.12.002

Thomas, A. W., Molter, F., Krajbich, I., Heekeren, H. R., & Mohr, P. N. C. (2019). Gaze bias differences capture individual choice behaviour. *Nature Human Behaviour*, *3*(6), 625–635. https://doi.org/10.1038/s41562-019-0584-8

Tian, K., Markkula, G., Wei, C., Sadraei, E., Hirose, T., Merat, N., & Romano, R. (2022). Impacts of visual and cognitive distractions and time pressure on pedestrian crossing behaviour: A simulator study. *Accident Analysis and Prevention*, *174*. https://doi.org/10.1016/j.aap.2022.106770

Tian, K., Tzigieras, A., Wei, C., Lee, Y. M., Holmes, C., Leonetti, M., Merat, N., Romano, R., & Markkula, G. (2023). Deceleration parameters as implicit communication signals for pedestrians' crossing decisions and estimations of automated vehicle behaviour. *Accident Analysis and Prevention*, *190*. https://doi.org/10.1016/j.aap.2023.107173

Tong, Y., Jia, B., & Bao, S. (2021). An augmented warning system for pedestrians: User interface design and algorithm development. *Applied Sciences (Switzerland)*, *11*(16). https://doi.org/10.3390/app11167197

Tran, T. T. M., Parker, C., Hoggenmüller, M., Hespanhol, L., & Tomitsch, M. (2023). Simulating Wearable Urban Augmented Reality Experiences in VR: Lessons Learnt from Designing Two Future Urban Interfaces. *Multimodal Technologies and Interaction*, *7*(2). https://doi.org/10.3390/mti7020021

Vinkhuyzen, E., & Cefkin, M. (2016). Developing Socially Acceptable Autonomous Vehicles. *Ethnographic Praxis in Industry Conference Proceedings*, *2016*(1), 522–534. https://doi.org/10.1111/1559-8918.2016.01108

Wilbrink, M., Nuttelmann, M., & Oehl, M. (2021). Scaling up automated vehicles ehmi communication designs to interactions with multiple pedestrians-putting ehmis to the test. *Adjunct Proceedings - 13th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2021*, 119–122. https://doi.org/10.1145/3473682.3480277

Yang, Y., Lee, Y. M., Madigan, R., Solernou, A., & Merat, N. (2024). Interpreting pedestrians' head movements when encountering automated vehicles at a virtual crossroad. *Transportation Research Part F: Traffic Psychology and Behaviour*, *103*, 340–352. https://doi.org/https://doi.org/10.1016/j.trf.2024.04.022