Does another pedestrian matter? -A Virtual Reality study on the interaction between multiple pedestrians and autonomous vehicles in shared space

Preprint · September 2023

DOI: 10.31219/osf.io/r3udx

Yan Feng
Delft University of Technology

Haneen Farah
Delft University of Technology

B. van Arem
Delft University of Technology

Does another pedestrian matter? – A Virtual Reality study on the interaction between multiple pedestrians and autonomous vehicles in shared space

YAN FENG[*]

Delft University of Technology, y.feng@tudelft.nl

ZHENLIN XU

Delft University of Technology, G.Xu-2@tudelft.nl

HANEEN FARAH

Delft University of Technology, H.Farah@tudelft.nl

BART VAN AREM

Delft University of Technology, B.vanArem@tudelft.nl

This study utilized Virtual Reality (VR) experiments to investigate pedestrian-autonomous vehicle interaction in shared spaces. In the VR experiment, pedestrians attempt to cross the road under different conditions, including the presence of another pedestrian, different external Human-Machin-Interfaces, AV driving styles, and road conditions. We employed an innovative VR setup that enabled two pedestrians to interact in real time with physical movements within an immersive VR environment. Overall, we found that the presence of multiple pedestrians significantly influenced pedestrian movement dynamics during road crossing. Additionally, the relative standing position had a significant impact on the distant pedestrians regarding time before crossing and vehicle-gazing behavior. While previous studies predominantly focused on pedestrian-AV interaction with a single pedestrian, this study takes an important step forward in terms of theory, methods, and relevance by considering interactions between multiple pedestrians and AVs. The findings establish a basis for further exploration of pedestrian-AV interaction in shared space.

CCS CONCEPTS • Human-centered computing → Human computer interaction (HCI) → Interaction paradigms → Virtual reality

Keywords: shared space, autonomous vehicles, multi pedestrians, VRU-AV interaction, eHMI, road crossing; VR

1 INTRODUCTION

Shared space is an urban area where pedestrians, cyclists, and vehicles are present without imposed traffic rules. It has become a popular urban planning approach to encourage low motorized traffic and create an urban space that is more accessible, safe, and social. With the rapid development of Autonomous Vehicles (AVs), it is expected that in the near future, more AVs will be employed on urban roads increasing the chances of their interaction with pedestrians. Thus it is crucial to understand the interaction between pedestrians and AVs in order to ensure pedestrians' safety as well as efficient operation of the AVs.

[*] Corresponding author.

Numerous studies investigated pedestrian-AV interaction with various focuses, including the effect of external human-machine interfaces (eHMIs) [1–3], AV's driving style [4–6], and road conditions [7–9]. It is challenging to study the interaction between pedestrians and AV due to safety, ethical, and financial constraints. Virtual Reality (VR) provides the possibility to study pedestrian-AV interaction in a safer setting with high experimental control, flexibility in modifying traffic scenarios, high accuracy of collected data, and acceptable ecological validity [10,11]. There is an increased number of studies that have employed VR to study pedestrian-AV interaction [7,12–14]. However, most studies simplify pedestrian-AV interaction by focusing on only a single pedestrian crossing in front of an AV. Crossing situations would be substantially more complex in reality. Moreover, the majority of studies focused on relatively traditional road scenarios (e.g., single-lane road crossing), while shared space introduces pedestrians to increased uncertainty and versatility because there are reduced or no traffic signs or road markings [15]. However, research shows that the decision-making process of road users is closely related to the complexity of the traffic environment [7]. To date, only a few studies investigated pedestrian-AV in shared spaces, focusing on interactions between a pedestrian and AV or relying on subjective questionnaire responses [8,16]. In summary, there is not much research into the interaction between multiple pedestrians and AVs in shared spaces, hampering a comprehensive understanding of their dynamic interactions.

To address this research gap, the current study employed VR experiments to investigate pedestrian-AV interaction in shared spaces in diverse conditions. These conditions include the presence of multiple pedestrians, different designs of eHMIs, different AV driving styles, and distinct road conditions. During the VR experiments, participants were able to physically cross the road and both objective (e.g., movement trajectory, gaze point) and subjective (e.g., user experience, trust in AV) data were collected. The impact of the above-mentioned factors on pedestrian road crossing behavior was analyzed using various behavioral metrics (e.g., crossing initiation time, gazing time, crossing speed). Moreover, we assessed the feasibility of employing a combination of multiple-user, real-walking locomotion, and an immersive VR system to study pedestrian-AV interaction.

The contribution of this paper is threefold. Firstly, this study presents a novel contribution to the body of work by investigating the interaction between multiple pedestrians and AV in shared spaces. Secondly, a variety of behavioral measures are analyzed to provide a holistic perspective of AV–pedestrian interactions. Thirdly, this study showcases the feasibility of incorporating multiple road users with an immersive VR setting for the examination of their interactions in more complex traffic scenarios.

The paper is organized as follows: section 2 presents studies that investigated pedestrian-AV interaction and the current research gap. Section 3 details the experiment method and modeling process. Accordingly, sections 4 and 5 present and discuss the results. The paper ends with a conclusion and future research directions.

2 BACKGROUND

The interaction between AVs and pedestrians has been investigated via a variety of methods, including questionnaire studies [3,8], controlled studies in real-world environments [17,18], and controlled studies in virtual reality environments [7,12–14]. These studies have focused on understanding the impact of eHMIs, AV driving styles, and traffic situations on pedestrian-AV interactions. This section gives an overview of the above-mentioned studies.

Informal communication between pedestrians and drivers is an essential input for pedestrians when making road-crossing decisions (e.g., gestures, eye contact) [19]. Given the reduced involvement or potential absence of drivers with AVs, it is important to understand the mechanisms of communication between pedestrians and AVs in order to ensure safe interaction between the two. Recently, significant research has been conducted to explore the role of eHMIs in increasing the efficiency of interactions between AVs and pedestrians [9,12,18,20,21]. Various eHMIs with diverse designs, display options, and technologies have been developed and tested. One group of studies focused on different forms of eHMIs, including text messages, lighting signals, animations, and more. Text messages normally convey the information regarding AV's status or advice to pedestrians. AV's status includes whether the AV is in automated modes [18,22] and AV's intention, such as "CAR STOPS" or "BRAKING" [23]. AVs can also give pedestrians direct advice to cross the road via text messages such as "WALK or "STOP" [22,24]. While some studies found this type of eHMI is helpful, other studies argue that it is more suitable for traffic situations involving a single pedestrian since it might lead to confusion when multiple road users are present [9]. Besides text messages, the awareness of pedestrians of AV's intention can be communicated via lighting conditions displayed on the vehicle as a combination of light color (e.g., red, green, blue) and lighting modes (e.g., static and dynamic) [18,24–26]. However, it is also found that eHMIs with lighting signals are less intuitive to pedestrians, and that often prior explanation is required [18,24]. In order to improve intuitive comprehensibility, some studies employed animated visuals with light patterns to convey information [3]. For instance, Othersen et al. (2018) [27] found that a walking-man animation had the best understandability and perceptibility compared to static lighting eHMI designs, similar to the finding of [12]. In addition to displaying or attaching the eHMIs on the vehicle, another group of studies explored the usage of projection-based eHMIs to convey AV's awareness of the pedestrian (i.e., that the AV detected the pedestrian). It is argued in these studies that projection-based eHMIs can be visible to multiple pedestrians and they utilize physical elements that are already embedded in the surroundings, such as pavements [22,28,29], road signs, traffic signals, and other infrastructure components.

In addition to investigating the impact of eHMIs on the efficient communication between pedestrians and AVs, several studies focused on implicit communication, specifically analyzing the state of vehicles. For instance, studies showed pedestrian crossing behavior can be influenced by a vehicle's kinematics information, such as vehicle speed [9] and time gap [5,13,14,30]. While some studies argue that implicit communication cues are more effective and efficient, a few studies showed that pedestrians rely on more vehicle kinematics to make crossing decisions [6,28,31]. Moreover, the acceleration and deceleration behavior of vehicles can also indicate their intention of yielding or not yielding, which have been found to affect pedestrians' crossing decisions [5,17,22].

The interaction between pedestrians and AVs has been investigated in diverse road-crossing situations. The majority of studies explored the interaction in unsignalized traffic situations, mostly featuring one-lane road [2,3,5] and two-lane road [30,32]. Compared to more traditional unsignalized traffic situations, a few studies investigated pedestrian-AV interaction in unmarked crossings, junctions, and intersections [5]. Studies indicate that pedestrian road crossing decision is related to the intricacy of traffic situations [7]. While pedestrian-AV interaction in various types of unsignalized traffic situations has been studied, their interaction in shared spaces has been rarely studied. The limited existing studies mainly depend on subjective questionnaire responses [8].

Literature suggests that other pedestrian's behavior is an important factor in pedestrian road-crossing decisions [33]. Even pedestrians who do not necessarily travel as a group together, they can be influenced by seeing someone crossing the road and as a result, modify their crossing decision [33,34]. Although crossing the street with other pedestrians is a common occurrence in real-world traffic situations, there have been limited studies investigating pedestrian-AV interaction when multiple road users are present. For example, [2] conducted a VR experiment to evaluate the impact of different eHMIs on pedestrian's willingness to cross when two pedestrians were present. They concluded that clear and unambiguous communication via eHMI is crucial in situations with multiple pedestrians. In an experiment setting with traditional vehicles, [35] investigated pedestrian road crossing behavior with a risky or safe computer-generated pedestrian, [32] and [36] asked one participant to cross the road alongside a group of pedestrians in a virtual environment. However, when there are more pedestrians involved in the above-mentioned studies, they are pre-programmed computer agents, or the type of interaction is limited (e.g., clicking a button to indicate the cross-decision). Hence, it's uncertain whether pedestrian road-crossing behavior would be the same if they were engaging with real humans in the traffic scenario [30]. Only [37] and [38] investigated paired participants crossing the road in a physical space together where the virtual environment was projected on large-screen displays. Both studies found that two pedestrians often crossed the same traffic gap together, although they were not instructed to do so. However, both studies only investigated pedestrian road crossing in traditional traffic situations without AVs.

To summarize, most of the above-mentioned studies investigated one-on-one pedestrian-AV interaction in relatively simple traffic scenarios, while shared spaces—a type of traffic situation characterized by increased uncertainty and versatility—have received less attention. Literature suggests that when the complexity of traffic situations increases, pedestrians perceive a higher risk and therefore behave more cautiously [7]. Moreover, some studies suggest that the scalability and suitability of eHMI need to be examined for more complex traffic scenarios [2,39]. Therefore, there is a need to investigate pedestrian-AV interaction in shared spaces with multiple pedestrians, aiming to understand the impact of eHMI, driving style, and road type on pedestrian crossing behavior in this type of traffic environment.

3 METHOD

3.1 Experiment design

The current study employed immersive Virtual Reality experiments to examine how various factors such as multiple pedestrians, eHMI, AV's driving style, and road conditions influence pedestrian road crossing behavior. A within-subject design approach was used in the current study to remove the effects of individual differences.

3.1.1Experiment scenario design

One existing shared space, namely the Marineterrein area in Amsterdam, the Netherlands, was chosen as a baseline to construct the VR environment. In this environment, there were no traffic lights, no stop signs, no pedestrian zebra, or any other elements to indicate the right of way. An audio soundscape was added to the VR environment to enhance the realism of the VR experience. Figure 1 shows the top view of the virtual environment. Three within-subject variables were included in the experiment, namely the number of pedestrians (i.e., 1, 2), type of path (i.e., straight path, T-junction path), type of eHMI (i.e., none, pedestrian sign, projected zebra), and AV's deceleration style (i.e., type I and type II). A detailed description of each variable is explained below.

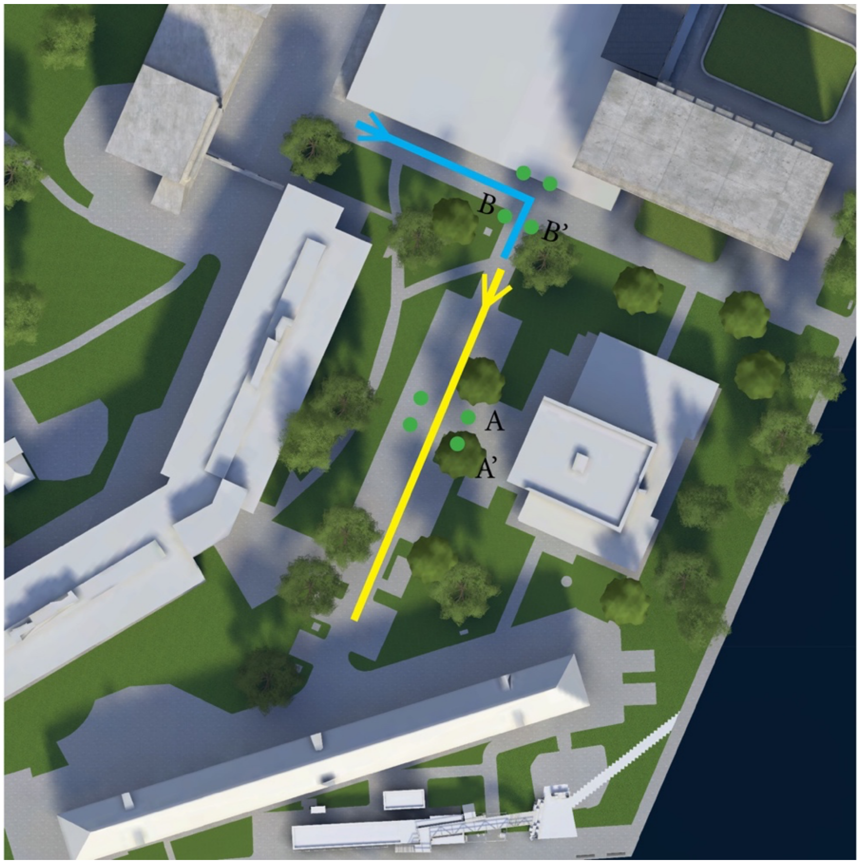

Figure 1. The top view of the virtual environment (the blue line is the T-junction path, the yellow line is the straight path, and green circles are the start and end positions of pedestrians).

Type of path: According to the study objective, two main paths and their surrounding areas were chosen as the experimental environment, including one straight path (yellow line in Figure 1) and one T-junction path (blue line in Figure 1). In all scenarios, the AV drove on the predefined path and operated according to specified driving behavior, mimicking the desired vehicle behavior of an AV. The arrows in Figure 1 indicate the approaching direction of the AV on both paths. For the T-junction path scenario, the AV indicated its turning intention by showing the turning light on. In the straight path scenarios, the AV approached the pedestrians from the right side of the participants; and in the T-junction path scenario, the AV approached the pedestrians from the left side of the participants.

Type of deceleration styles: Two types of deceleration styles were designed. For type I deceleration, when the distance between the vehicle and the pedestrian was 15 meters, the AV started to slow down from 15 km/h to 10 km/h with a deceleration rate of 2.5 m/s^2 and continued driving at 10km/h. When the AV was 5 meters away from the pedestrian, it started to reduce its speed to 5km/h and kept that speed until the distance was 3 meters. Finally, the AV stops moving when the distance between the pedestrian and the AV was equal to 3 meters. For type II deceleration, only one phase of deceleration was involved. When the distance between the vehicle and the pedestrian was 15 meters, the AV started to slow down from 15 km/h to 5 km/h with a deceleration rate of 2.5 m/s^2 and continued driving at 5 km/h. The AV stopped moving at a distance of 3 meters from the pedestrians. Type II deceleration was designed more defensive than Type I deceleration. Figure 2 illustrates the relationship between time to collision and the relative distance between the pedestrian and the AV, categorized by AV's deceleration type.

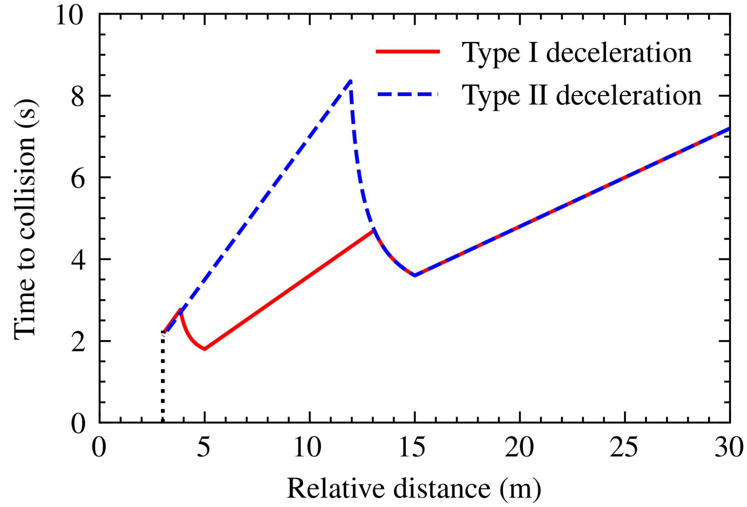

Figure 2. Relationship between time to collision and the relative distance between the AV and the pedestrian by deceleration type.

Number of pedestrians: For the single-pedestrian scenario, only one participant was immersed in the virtual environment to perform the road crossing task. One participant initially appeared at a pre-defined location, i.e., location A in the straight path scenario and location B in the T-junction path scenario (see Figure 1). For the two-pedestrian scenario, two participants were immersed and they could see each other in the virtual environment. The participant's body is represented by a virtual avatar with a head and shoulder, and their movement in real life was synchronized to the avatar in the virtual environment. Both participants need to perform the road crossing task at locations A and A' (straight path scenario), or B and B' (T-junction path scenario). The initial distance between two pedestrians was 4 meters, which was selected to strike a balance of not being too close to encourage group behavior, yet close enough that the yielding message from the eHMI could apply to both pedestrians.

Type of eHMI: To notify pedestrians about the yielding intention of the AV, two types of eHMI design were chosen for further investigation in the current study. Three levels of eHMI were included, namely none eHMI, eHMI with a pedestrian sign on the AV's windshield, and eHMI with a projected zebra on the road. Figure 3 shows the overview of the eHMI tested in the current study. For the pedestrian sign eHMI (Figure 3b), a static green-color pedestrian sign was displayed in the middle of the windshield to indicate that the AV intends to yield to the pedestrian and that the pedestrian can cross the road. Another eHMI concept was adapted based on the design of [40] and [29]. When the AV started to yield, the AV projected a green-color crossing zebra on the road to indicate that the vehicle intends to yield to the pedestrian (Figure 3c). These eHMIs were activated when the distance between the (first) pedestrian and the AV was equal to 5.6 meters.

a. AV without eHMI b. AV with pedestrian-sign c. AV with projected zebra

Figure 3. The overview of tested eHMIs

3.1.2Experiment task design

The combination of all variables resulted in a total of 18 road-crossing scenarios. These scenarios were further divided into two blocks. The first block includes 12 single-pedestrian scenarios (2 path x 2 deceleration x 3 eHMIs) and the second block includes 6 two-pedestrian scenarios (2 path x 3 eHMIs). Only the within-subject variables differ among different scenarios, the rest of the infrastructure, including the surrounding buildings, sounds, etc. remained the same. Each participant encountered the first block and then the second block. The scenarios within each block were randomized to reduce learning effects. For the two-pedestrian scenario, the relative standing position of the two participants was randomly assigned, either closer or further away from the AV.

At the beginning of each scenario, the participant stood facing the corresponding street. There was a green circle on the ground near the curbside that indicated the starting position of the road-crossing task (see Figure 1). When the participant stepped into the green circle from the initial position, the AV started to approach the pedestrian from 30 meters away at a speed of 15km/h in accordance with the speed limit of shared space in the Netherlands. At the same time, another green circle was activated and appeared on the opposite side of the road to indicate the ending position (see Figure 1). Participants were instructed to cross the street at the last moment they feel safe to do so. Their task was described as follows: '*Please cross at the last moment you feel safe to cross*'. The instruction design was adopted from the study of [41].

3.2 Experiment Apparatus

The VR experiment was conducted in a room that is 15 meters long x 8 meters wide x 3.8 meters high. The room was divided into two parts equally for the single-pedestrian scenarios and used together for the two-pedestrian scenarios. HTC VIVE Pro Eye headset (resolution: 1440 x 1600 pixels per eye, 110 of field-of-view, a 90HZ refresh rate) and HP Reverb G2 Omnicept headset (resolution: 2160 x 2160 pixels per eye, 114 of field-of-view, a 90HZ refresh rate) were used during the experiment. The HTC headset was connected to a Windows 10 desktop that was equipped with an Intel(R) Core (TM) i7-8700 CPU, a 16 GB RAM, an NVIDIA GeForce RTX 2070 graphics card, and a SanDisk SD9SN8W 256 GB SSD. The HR headset was run on a Windows 10 desktop that was based on an Intel(R) Core (TM) i7-10700F CPU, a 16 GB 2933Mhz of RAM, and an NVIDIA GeForce RTX 3060 GPU. Figure 4 illustrates two persons wearing the two headsets in the experimental setting.

Figure 4. Illustration of two persons wearing two Virtual Reality headsets.

In the current study, we utilized the real-walking locomotion style, allowing users to have continuous motion regarding movements and rotations in the real world, which can be matched under a 1:1 scheme to the virtual environment. Participants were able to move in the real-life environment under a 1:1 scheme mapped to the virtual environment. Literature shows that real-walking locomotion enables more realistic, natural walking movement, and leads to a higher sense of presence compared to other locomotion styles [42–44].

3.3 Experiment procedure

The experiment procedure includes four parts, namely (1) introduction of the experiment, (2) familiarization with the VR system, (3) official experiment, and (4) filling in the post-questionnaire. Ethical approval was obtained from the Human Research Ethics Committee of the Delft University of Technology (Reference ID: 2042). These four parts are further elaborated:

1. Introduction: This part includes providing participants with written information about the experiment procedure, the explanation of AV, the demonstration, and the meaning of the tested eHMIs. They were also informed that they could stop the experiment at any time if they felt uncomfortable. Accordingly, participants read and signed the consent form.

2. Familiarization: Participants were invited to wear the headset and adjust the headphone properly. Then they were immersed in a simple virtual environment to walk around in order to familiarize themselves with the locomotion method. There were no other pedestrians or vehicles presented.

3. VR experiment: After the familiarization part, participants were instructed to stand at a predefined location marked with black tape in the experiment room and face in the right direction. Then they were randomly assigned to one experiment scenario of the first block and asked to perform the experiment tasks. After each experiment scenario, participants had to walk back to their predefined location before the next experiment scenario started. After the first block (i.e., single-pedestrian scenario) was completed, two participants were teleported to the second block (i.e., two-pedestrian scenario). Two participants were asked to move to the first green circle with a three-second countdown by the experimenter.

4. Filling in the post-questionnaire: After finishing the VR experiment, participants were asked to remove the headset. Then they were asked to fill in a post-questionnaire in the same experiment room. Afterward, participants were thanked and received around a € 20 voucher as compensation.

3.4 Data collection

Two types of data were collected during the experiment, including objective data (i.e., movement trajectory, gaze point) and subjective data (i.e., questionnaire data).

Within Unity, the participant's movement in the virtual environment was recorded. The data recording started when participants arrived at the first green circle and ended when participants reached the second green circle. The collected data included (1) timestamp, (2) participant's position (i.e., coordinate x, y, z), (3) head rotation (i.e., roll, yaw, pitch), and (4) gaze point (i.e., coordinate x, y, z). All data were recorded at a frequency of 20 HZ.

The questionnaire included seven parts, namely (1) participant's information, (2) the face validity questionnaire, (3) the Simulator Sickness Questionnaire, (4) the Presence Questionnaire, (5) the Trust in AVs questionnaire, (6) the Perceived behavioral control and risk questionnaire, and (7) the System Usability Scale questionnaire. The personal information part included participants' characteristics such as gender, age, achieved highest education level, familiarity with the Marineterrein area, familiarity with computer gaming, familiarity with VR, familiarity with the concept of AVs, and experience regarding interaction with AVs. The face validity questionnaire measured whether a simulator measures what it is intended to measure [45]. Within the face validity questionnaire, the realism of the virtual environment, virtual objects (e.g., the vehicle), movement ability, and environmental sound were rated on a 5-point scale. The Simulator Sickness Questionnaire is a standard questionnaire [46] to measure the experienced simulation sickness of participants in the virtual environment. The Presence Questionnaire [47] measured the feeling of presence in the virtual environment. The Trust in AVs questionnaire was adopted based on the study of [48], which contained 7 items including questions such as 'During the experiment, I trust the automated vehicle to keep its distance from me.' And 'During the experiment, I trust the automated vehicle to drive safely.' The Perceived behavioral control (PBC) and Perceived Risk (PR) included 3 items, namely 'For me, crossing the road in this way would be …', 'I believe that I have the ability to cross the road in this way as described in this situation', and 'Crossing the road in the way as described in this situation would be…'. Finally, the System Usability Scale questionnaire [49] assessed the usability of the VR system.

3.5 Participant's characteristics

In total, 54 participants aged between 17 and 76 years old (*M* = 33.63, *SD* = 14.08) were recruited and took part in the experiment. All participants had normal vision or corrected vision and normal mobility. None of the participants dropped out of the experiment due to motion sickness. In the end, 50 participants finished both single-pedestrian scenarios and two-pedestrian scenarios; 4 participants didn't perform the two-pedestrian scenarios because there was not another participant present at the same time during the experiment. The characteristics of the participants are shown in Table 1.

Table 1: Demographic information of participants

Descriptive information	Category	Number (percentage)
Gender	Male	27 (50.00%)
	Female	27 (50.00%)
Highest education level	High school or equivalent	2 (3.70%)
	Associate degree or equivalent	3 (5.56%)
	Bachelor’s degree or equivalent	27 (50.00%)
	Master’s degree or equivalent	20 (37.04%)
	Doctoral degree or equivalent	2 (3.70%)
Previous experience with VR	Never	17 (31.48%)
	Seldom	25 (46.30%)
	Sometimes	9 (16.67%)
	Often	1 (1.85%)
	Very often	2 (3.70%)
Familiarity with any computer gaming	Not at all familiar	10 (18.52%)
	A-little familiar	12 (22.22%)
	Moderately familiar	14 (25.93%)
	Quite-a-bit familiar	6 (11.11%)
	Very familiar	12 (22.22%)
Familiarity with the Marineterrein area	Not at all familiar	4 (7.41%)
	A-little familiar	3 (5.56%)
	Moderately familiar	4 (7.41%)
	Quite-a-bit familiar	12 (22.22%)
	Very familiar	31 (57.41%)
Familiarity with the concept of automated shuttles	Not at all familiar	4 (7.41%)
	A-little familiar	9 (16.67%)
	Moderately familiar	16 (29.63%)
	Quite-a-bit familiar	16 (29.63%)
	Very familiar	9 (16.67%)
Previous experience with automated shuttles	Never	41 (75.93%)
	Seldom	11 (20.37%)
	Sometimes	1 (1.85%)
	Often	1 (1.85%)
	Very often	0 (0.00%)

3.6 Data analysis

Various indicators can be derived from the objective data (i.e., movement trajectory, gaze point) collected during the experiments. The indicators selected for data analysis in this study are defined as follows:

- Time Before Crossing (TBC): T_{TBC} refers to the duration during which the pedestrian spends waiting before initiating the crossing from the moment the experiment starts (i.e., from the moment the pedestrian stepped in the green circle).
- Crossing-initiation time (CIT): T_{CIT} is calculated as the period between the moment that the pedestrian sees the AV (extracted from the gaze point data) and the moment the pedestrian starts to cross. If the pedestrian notices the AV before initiating a crossing, then CIT is positive. Otherwise, CIT is negative.
- Time to Cross (TTC): T_{TTC} is defined as the duration it takes for the participant to reach the other side of the road from the moment they begin the road crossing.

- Vehicle-gaze time (T_{AV}): T_{AV} is aggregated by the collected eye-gazing data and means the total duration of gazing on the AV during the whole crossing process.
- Total crossing distance (D): is the total distance traveled by the pedestrian during road crossing.
- Crossing speed (v): is the mean crossing speed calculated by dividing total crossing distance by totoal crossing time.
- Space gap (L): is the distance between the AV and the pedestrian when the pedestrian starts to cross.

The linear mixed model (LMM) was employed to study the influence of several factors including multiple pedestrians, eHMI, AV driving style, and road conditions on pedestrian crossing behavior. Seven dependent variables, namely crossing initiation time T_{CIT}, time before crossing T_{TBC}, time to cross T_{TTC}, vehicle-gazing time T_{AV}, total crossing distance D, crossing speed v were modeled, respectively. The LMM is a function of spacing gap L, road type μ_{road}, eHMI type μ_{eHMI}, AV driving style μ_{AV}, the number of participants μ_{single}, position of participant $Position$, distance difference of two pedestrians ΔD^{2ped}. We modeled the independent variables and their interaction as fixed effects and the ID of pedestrians as random effects by the maximum likelihood estimation method. The random intercepts were only considered in the models in terms of the random effects. The model formulations of two families were defined as follows in Equation 1 using Wilkson notation, where $\mu_{single} * Position * \Delta D^{2ped}$ is an interaction factor. Among these, μ_{single} is a binary variable that indicates whether only one participant takes part in the experiment or if there are multiple participants involved. $Position1$ and $Position2$ are two dummy variables indicating the initial position of the two participants. Specifically, $Position1$ indicates the participant starts from the closer side to the vehicle side and $Position2$ means that the initial position is further away from the vehicle. Difference in distance ΔD^{2ped} is the distance between two participants when they decide to cross. μ_{road} is the binary variable to indicate the road type within the scenario is straight path ($\mu_{road} = 0$) or T-junction path ($\mu_{road} = 1$). μ_{eHMI1} and μ_{eHMI2} are two dummy variables to indicate the type of eHMI with the reference level meaning none eHMI is equipped with the AV. μ_{AV} is also another binary variable indicating the AV driving style set within the scenario. The R programming language (Version 4.2.3) and lmerTest library (Version 3.1-3) were used for the statistical modeling and analysis.

$$T_{CIT}/T_{TBC}/T_{TTC}/T_{AV}/D/v \sim L + \mu_{single} * Position * \Delta D^{2ped} + \mu_{road} + \mu_{eHMI1} + \mu_{eHMI2} + \mu_{AV} + (1|\#Participant) \quad (1)$$

4 RESULTS

4.1 Time before crossing (TBC)

As shown in Table 2, the space gap, Type II deceleration, and the distant participants were found to influence the TBC significantly. The space gap (β_L) exhibited a negative effect on the TBC, indicating that the pedestrians needed less time to initiate their crossing as the gap increased. Type II deceleration, in contrast to type I deceleration, caused pedestrians to take longer preparation time before-crossing. Moreover, the number of pedestrians involved in the crossing process and their relative positions had a significant impact on the TBC. In the two-pedestrian scenarios, the pedestrian who was further away from the AV spent an additional around 0.21 seconds to initiate their actions as the distance between the two participants increased by 1 meter at the moment they decide to start crossing.

Table 2: The results for linear mixed model of time before crossing (TBC)

Predictors	***Est.***	***SE***	***CI***	***p***

(Intercept)	7.206	0.101	[7.008, 7.404]	<0.001
$\boldsymbol{\beta_{L}}$	**-0.232**	**0.004**	**[-0.240, -0.225]**	**<0.001**
β_{road} (T-junction)	-0.006	0.019	[-0.030, 0.043]	0.713
β_{eHMI1} (pedestrian)	0.011	0.023	[-0.034, 0.056]	0.626
β_{eHMI2} (zebra)	0.019	0.023	[-0.025, 0.064]	0.398
$\boldsymbol{\beta_{AV}}$ **(Type II)**	**-0.072**	**0.022**	**[-0.114, -0.029]**	**0.001**
$\beta_{IsSingle*Position1*\Delta D^{2ped}}$	-9.896e-04	0.007	[-0.015, 0.013]	0.890
$\boldsymbol{\beta_{IsSingle*Position2*\Delta D^{2ped}}}$	**0.210**	**0.008**	**[0.193, 0.225]**	**<0.001**
Random effects	***Var***	***SD***	***p***	
$\#Participant$: Intercept	0.002	0.046	0.047	
Model Performance				
Observations	939			
Marginal R^2 / Conditional R^2	0.788 / 0.794			
logLik	-161.0			
AIC	342.1			
BIC	390.5			

4.2 Crossing initiation time (CIT)

Table 3 shows the results of the mixed-effects modeling with random intercepts for CIT. As shown in Table 3, the space gap had a significant negative effect on the CIT, implying that pedestrians were less hesitant to begin their crossing as the spacing gap was larger. Among the four main factors investigated in this research, road type was the only one that exhibited a significant impact on the CIT. Specifically, pedestrians tended to allocate a longer time (0.363 seconds more) to decide to cross the road at the T-junction path compared to the straight path. Moreover, it is interesting to note that both eHMIs do not significantly affect CIT in the current study.

Table 3: The results for linear mixed model of crossing initiation time (CIT)

Predictors	***Est.***	***SE***	***CI***	***p***
(Intercept)	5.818	0.636	[4.571, 7.064]	<0.001
$\boldsymbol{\beta_{L}}$	**-0.257**	**0.024**	**[-0.305, -0.209]**	**<0.001**
$\boldsymbol{\beta_{road}}$ **(T-junction)**	**0.363**	**0.112**	**[0.144, 0.582]**	**0.001**
β_{eHMI1} (pedestrian)	-0.008	0.137	[-0.277, 0.261]	0.953
β_{eHMI2} (zebra)	-0.023	0.137	[-0.291, 0.245]	0.865
β_{AV} (Type II)	-0.027	0.132	[-0.285, 0.231]	0.835
$\beta_{IsSingle*Position1*\Delta D^{2ped}}$	-0.002	0.044	[-0.088, 0.084]	0.962
$\beta_{IsSingle*Position2*\Delta D^{2ped}}$	0.033	0.049	[-0.064, 0.129]	0.506
Random effects	***Var***	***SD***	***p***	
$\#Participant$: Intercept	0.963	0.981	<0.001	

Model Performance	
Observations	939
Marginal R^2 / Conditional R^2	0.111 / 0.331
logLik	-1888.8
AIC	3797.5
BIC	3846.0

4.3 Time to cross (TTC)

Table 4 presents the results for linear mixed model of TTC. Among the factors examined, only the space gap demonstrated a statistically significant effect on the TTC. Specifically, pedestrians took longer to cross the road when the spacing gap between them and the AV was larger. On the other hand, the road type, eHMI design, AV driving style, the number of pedestrians and their relative positions did not show any significant impact on TTC. These results suggest that the space gap is a crucial factor influencing the crossing time for pedestrians, while other variables did not affect the TTC.

Table 4: The results for linear mixed model of time to cross (TTC)

Predictors	***Est.***	***SE***	***CI***	***p***
(Intercept)	3.599	1.575	[0.534, 6.686]	0.022
$\boldsymbol{\beta_L}$	**0.339**	**0.060**	**[0.217, 0.453]**	**<0.001**
β_{road} (T-junction)	-0.242	0.276	[-1.038, 0.237]	0.382
β_{eHMI1} (pedestrian)	-0.174	0.339	[-0.770, 0.828]	0.607
β_{eHMI2} (zebra)	-0.145	0.338	[-0.797, 0.750]	0.669
β_{AV} (Type II)	0.138	0.325	[-0.403, 0.872]	0.671
$\beta_{IsSingle*Position1*\Delta D^{2ped}}$	0.006	0.109	[-0.322, 0.166]	0.954
$\beta_{IsSingle*Position2*\Delta D^{2ped}}$	-0.040	0.121	[-0.389, 0.144]	0.743
Random effects	***Var***	***SD***	***p***	
$\#Participant$: Intercept	6.504	2.550	<0.001	
Model performance				
Observations	939			
Marginal R^2 / Conditional R^2	0.032 / 0.290			
logLik	-2739.5			
AIC	5499.0			
BIC	5547.4			

4.4 Vehicle-gazing time

Table 5 reveals the significant impact of the road condition, the type of eHMI, and the number of pedestrians on the vehicle-gaze time. The T-junction path had a positive effect on the vehicle-gaze time, meaning that pedestrians tended to spend more time observing the approaching vehicle both before and during the crossing phase at the T-junction path compared to the straight path. This could be because of the uncertainty involved with the turning direction of the AV approaching the T-junction. Regarding the eHMI design, both versions exhibited a reduction in vehicle-gazing time. This finding suggests that eHMI effectively provided pedestrians with quick and sufficient information to determine whether to proceed with their street crossing or not, in contrast to AVs without eHMI. It is also interesting to note that there was not much difference regarding vehicle-gazing time between the two eHMIs conditions. Regarding the number of pedestrians crossing and their relative positions, the pedestrians close to the AV tended to spend more time observing the AV but the increase was not statistically significant. The pedestrian who was further away from the AV tended to significantly reduce their gazing time on the AV. Moreover, we also derived the gazing time towards another person in the two-pedestrian scenarios before they decide to cross. The results showed that participants who were further away from the AV spent more time looking at another pedestrian (M = 0.18, SD = 0.33) compared to the participants who were closer to the AV (M = 0.06, SD = 0.17).

Table 5: The results for linear mixed model of vehicle-gazing time

Predictors	***Est.***	***SE***	***CI***	***p***
(Intercept)	4.584	0.935	[2.751, 6.416]	<0.001
β_L	0.001	0.035	[-0.067, 0.070]	0.974
$\boldsymbol{\beta_{road}}$ (T-junction)	**0.373**	**0.160**	**[0.059, 0.688]**	**0.020**
$\boldsymbol{\beta_{eHMI1}}$ (pedestrian)	**-0.730**	**0.197**	**[-1.115, -0.344]**	**0.002**
$\boldsymbol{\beta_{eHMI2}}$ (zebra)	**-0.710**	**0.196**	**[-1.095, -0.326]**	**<0.001**
β_{AV} (Type II)	0.016	0.189	[-0.355, 0.386]	0.934
$\beta_{IsSingle*Position1*\Delta D^{2ped}}$	0.087	0.063	[-0.037, 0.212]	0.168
$\boldsymbol{\beta_{IsSingle*Position2*\Delta D^{2ped}}}$	**-0.210**	**0.071**	**[-0.349, -0.072]**	**0.003**
Random effects	***Var***	***SD***	***p***	
$\#Participant$: Intercept	3.959	1.990	<0.001	
Model Performance				
Observations	939			
Marginal R^2 / Conditional R^2	0.026 / 0.412			
logLik	-2244.2			
AIC	4508.4			
BIC	4556.8			

4.5 Total crossing distance

Table 6 demonstrates that apart from the space gap, the number of participants, and their relative position had significant impact on the total crossing distance. The space gap showed a slight positive effect on the total distance pedestrians walked during crossing. In the two-pedestrian scenarios, both participants had a shorter total crossing distance, indicating that they crossed the road following the shortest path.

Table 6: The results for linear mixed model of total crossing distance

Predictors	***Est.***	***SE***	***CI***	***p***
(Intercept)	3.177	0.263	[2.661, 3.694]	<0.001

$\boldsymbol{\beta_L}$	**0.066**	**0.010**	**[0.046, 0.086]**	**<0.001**
β_{road} (T-junction)	0.081	0.047	[-0.011, 0.174]	0.086
β_{eHMI1} (pedestrian)	-0.019	0.058	[-0.130, 0.098]	0.784
β_{eHMI2} (zebra)	0.024	0.058	[-0.090, 0.137]	0.682
β_{AV} (Type II)	-0.082	0.056	[-0.191, 0.027]	0.140
$\boldsymbol{\beta_{IsSingle*Position1*\Delta D^{2ped}}}$	**-0.074**	**0.019**	**[-0.110, -0.038]**	**<0.001**
$\boldsymbol{\beta_{IsSingle*Position2*\Delta D^{2ped}}}$	**-0.142**	**0.021**	**[-0.182, -0.102]**	**<0.001**
Random effects	***Var***	***SD***	***p***	
$\#Participant$: Intercept	0.069	0.263	<0.001	
Model Performance				
Observations	939			
Marginal R^2 / Conditional R^2	0.069 / 0.178			
logLik	-1060.7			
AIC	2141.3			
BIC	2189.8			

4.6 Crossing Speed

The results of linear mixed model for crossing speed are shown in Table 7. The crossing speed is significantly affected by the space gap, road type, declaration type, and the number of participants. The space gap had a slightly positive impact on crossing speed. The T-junction increased the crossing speed by 0.019 m/s. This could be because in more complex situations pedestrians were more aware of risk. Type II deceleration reduced the crossing speed by 0.014 m/s. This is according to expectations as Type II driving style is more defensive. The relative positions between the two pedestrians crossing the street also played a significant role. As the distance increased, both pedestrians reduced their crossing speed, with the pedestrian farther from the AV exhibiting a particularly significant reduction.

Table 7. The results for linear mixed model of crossing speed

Predictors	***Est.***	***SE***	***CI***	***p***
(Intercept)	0.292	0.027	[0.239, 0.345]	<0.001
$\boldsymbol{\beta_L}$	**0.004**	**0.001**	**[0.002, 0.006]**	**<0.001**
$\boldsymbol{\beta_{road}}$ **(T-junction)**	**0.019**	**0.005**	**[0.010, 0.028]**	**<0.001**
β_{eHMI1} (Windshield sign)	0.003	0.006	[-0.008, 0.014]	0.611
β_{eHMI2} (Protected zebra)	0.007	0.006	[-0.004, 0.018]	0.220
$\boldsymbol{\beta_{AV}}$ **(Type II)**	**-0.014**	**0.005**	**[-0.025, -0.004]**	**0.009**
$\boldsymbol{\beta_{IsSingle*Position1*\Delta D^{2ped}}}$	**-0.009**	**0.002**	**[-0.013, -0.006]**	**<0.001**
$\boldsymbol{\beta_{IsSingle*Position2*\Delta D^{2ped}}}$	**-0.020**	**0.002**	**[-0.024, -0.016]**	**<0.001**
Random effects	***Var***	***SD***	***p***	
$\#Participant$: Intercept	0.003	0.053	<0.001	

Model Performance	
Observations	939
Marginal R^2 / Conditional R^2	0.083 / 0.406
logLik	1083.5
AIC	-2147.0
BIC	-2098.6

4.7 Subjective measures

4.7.1 Realism

Using the face validity questionnaire, the realism of the participants' experience in the VR scenarios was rated based on the realism of the virtual environment, virtual objectives, movement ability, and environmental sound as displayed in Table 8. Among four items, the realism of movement ability received the highest score (*M* = 3.78, *SD* = 0.84), which shows that participants were able to move and walk in the virtual environment in a realistic manner using the adopted real-walking location style. The realism of the virtual environment received the lowest score (*M* = 3.48, *SD* = 0.79). Given that most participants were well acquainted with the experimental location in reality, it is plausible that they had higher expectations for the authenticity of the virtual environment. The average score of the face validity questionnaire is 3.68 (*SD* = 0.57). Similar scores were found in prior studies that employed VR to study pedestrian road crossing behavior [50]. During the initial encounter with the AV, we even observed one participant feeling uncertain about the AV's action and decided to run to the opposite side of the road. In general, both participants' ratings and researchers' observations validated the realism provided by the current VR setup.

Table 8: Rating of realism (range from 1 to 5)

Items	**Mean**	**SD**
The realism of the virtual environment	3.48	0.79
The realism of the virtual objects (e.g., automated shuttles)	3.68	0.67
The realism of the movement ability	3.78	0.84
The realism of the environmental sound	3.76	1.04

4.7.2Simulation sickness

To measure the extent of simulation sickness participants experienced in the virtual environment, we employed the well-established Simulator Sickness Questionnaire [46]. Using a 4-point Likert scale ranging from 0 (none) to 3 (severe), participants assessed 16 possible symptoms (e.g., eyestrain, nausea, vertigo). The ratings are categorized into three subscales representing symptoms of nausea, oculomotor disturbance, and disorientation. To obtain the subscale scores, the reported scores for each symptom were multiplied by their respective weights for that particular subscale. The result of each subscale is presented in Table 9, highlighting that Nausea received the lowest score and disorientation received the highest score. The relatively high score of disorientation could be attributed to participants needing to turn and return to their original position after completing each experimental scenario. The average score of the total SSQ is 28.40 (*SD* = 23.23), and no participants reported any discomfort or notable symptoms. According to both the SSQ results and participant feedback, the current study only revealed no or slight symptoms.

Table 9: Subscales of SSQ: Means and standard deviations

Subscale	Mean	SD
Nausea	15.55	16.50
Oculomotor disturbance	25.41	21.19
Disorientation	36.35	34.88

4.7.3Feeling of presence

Assessing the feeling of presence is crucial to ensure participants experience an engaging and immersive experience during the VR experiment. The current study employed the Presence Questionnaire (PQ) [47] to measure participants' sense of presence, which consists of four subscales including involvement, sensor fidelity, immersion, and interface quality. Participants rated 29 items using a 7-point Likert scale. The results of the PQ questionnaire are presented in Table 10. The highest score on the Immersion subscale indicates that the participants experienced a strong sense of immersion. The average total score of PQ in this study is 134.96 (*SD* = 19.25), indicating a strong sense of presence in the current study.

Table 10: Subscales of PQ (range from 1 to 7)

	Involvement	Sensory fidelity	Immersion	Interface quality [a]
Mean	4.77	3.80	5.38	3.96
SD	0.81	0.83	0.76	1.10

[a] Reversed items

4.7.4Trust in AVs

The level of trust in AVs was measured per participant using the a scale ranging from 1 to 7 [48,51]. This scale contained questions such as 'Globally, I trust the automated vehicle', 'I trust the automated vehicle to have seen me', and 'I trust the automated vehicle to drive safe'. In the current experiment, the mean score was 4.42 (*SD* = 1.09), indicating a moderate level of trust in the AV.

4.7.5Perceived Behavioral Control and Risk

The Perceived Behavioral Control (PBC) questionnaire was measured by 2 items, namely 'For me, crossing the road in this way would be …', and 'I believe that I have the ability to cross the road in this way as described in this situation'. The mean score of PBC is 5.63 (*SD* = 0.96). For the Perceived Risk (PR) questionnaire, participants answered the question 'Crossing the road in the way as described in this situation would be…' on a scale from 1 (very unsafe) to 7 (very safe). The mean score of PR is 5.09 (*SD* = 1.15). Both results of PBC and PR are similar to the study of [52].

4.7.6Usability

To ensure the usability of the VR setup for participants, the System Usability Scale (SUS) questionnaire was adopted [49]. It contains 10 items that participants rated from strongly disagree (1) to strongly agree (5). The total score of SUS can range from 0 to 100. In this study, the average score of SUS is 72.04 (*SD* = 13.30), which suggests 'good' usability based on the interpretation by [53].

5 DISCUSSIONS

In this study, we investigated the pedestrian-AV interaction in shared spaces in various conditions. Using the objective data collected via VR, our study specifically examines the impact of the presence of multiple pedestrians, different designs of eHMIs, different AV driving styles, distinct road conditions, and space gaps on road crossing behavior. Meanwhile, user experience was analyzed using the subjective data collected via the post-experiment questionnaire.

The results show that the pedestrian's relative standing position to the AV and the distance difference between them when they decided to cross the road had a significant impact on the time before crossing, gazing time towards the vehicle, total crossing distance, and crossing speed. For both pedestrians, their crossing distance and crossing speed were significantly reduced when two pedestrians were presented in the virtual environment. Studies suggested that when pedestrians are aware of other people being present in the same environment, they would reduce speed to avoid collisions [36]. We also found that pedestrians who were further away from the AV had longer decision time to cross the road and shorter gazing time towards the AV. Our interpretation is that when the far-away participant experienced the second block (i.e., two-pedestrian scenarios) after the first block (single-pedestrian scenarios), the sudden appearance of another person in the environment distracted and caught their attention, leading to longer decision time. Studies suggest that distracted pedestrians tend to initiate late road crossing [7], this is further confirmed by the longer gazing time toward another pedestrian. This finding is in line with [38], which similarly observed that participants farther from the AV chose narrower time margins to enter the road. Overall, our findings confirmed that neighbor's behavior can impact the pedestrian's movement dynamic when crossing the road [35]. Moreover, our findings suggest that when crossing the road next to each other, two pedestrians could behave differently depending on their relative standing positions.

Regarding the impact of eHMI on pedestrian crossing behavior, we found that both pedestrian-eHMI and zebra-eHMI only had a significant impact on vehicle-gazing time. There was no effect of eHMIs on pedestrian's decision to cross the road. This finding contrasts with literature suggesting the positive impact of eHMIs on pedestrian crossing decisions [9,12,24]. One possible reason is that in our study, compared to other studies, AVs operated at slower speeds in a shared space, resulting in participants encountering no critical or unexpected situations. Moreover, we found that participants had less observation time at the AV when it was equipped with eHMIs. Our finding suggests that eHMIs in shared spaces play a substantial role in guiding pedestrians' attention toward the AV. However, it may not be beneficial in reducing the decision-making time for road crossing.

In terms of the impact of AV's driving style, we found it has a significant impact on time before crossing, total crossing distance, and crossing speed. Our finding shows that more defensive deceleration behavior exhibited by the AV (i.e., type II deceleration) had a positive effect on the crossing decision, namely, participants took a shorter time before deciding to cross the road. There are two possible explanations for this. Firstly, as shown in Figure 2, when the distance between pedestrians and the AV was less than approximately 13 meters, the time to collision was higher in Type II declaration. It implies that participants perceived a greater level of safety to cross, resulting in shorter decision time. Second, a more defensive deceleration results in an earlier reduction in driving speed, which can better indicate the yielding behavior of a vehicle. This is also reflected by the results of slower crossing speed during type II deceleration. This finding is in line with other studies that suggest that early braking can better reflect AV's yielding intention and thus reduce pedestrian decision time to cross the road [14,17,54].

Regarding the impact of road type, we found that T-junction had a significant impact on crossing initiation time, vehicle-gazing time towards the vehicle, and crossing speed. When pedestrians cross the T-junction path in front of an AV, they tend to initiate the crossing decision later due to the increased uncertainty regarding AV's driving direction, which is in line with [7]. Regarding crossing speed, our results are similar to [7] who recorded a significantly higher average speed and insignificantly longer observation time toward the AV at the T-junction. Our findings indicated that pedestrians were more cautious in more complex traffic scenarios.

Regarding the space gap, we found it has a significant impact on crossing initiation time, time before crossing, time to cross, total crossing distance, and crossing speed. In line with other studies, we found the increased space gap led to less time to initiate the crossing decision, indicating pedestrians prefer to cross when there are larger spatial traffic gaps [5,14,30]. Moreover, we found participants took longer time and longer distances to cross the road when the distance from the vehicle was larger. Meanwhile, participants chose to cross the street faster when the space gap was larger. This is contrary to our expectations because a higher space gap should provide pedestrians with a greater margin of safety for crossing. It remains unclear why participants in the current study chose to cross faster with a larger space gap.

Compared to other VR studies that employed one-to-one interaction [7,9,14], this study employed a unique VR system that incorporated multiple users and real-walking locomotion to study pedestrian-AV interaction. The results of user experience show that participants had a positive experience using this VR system, indicated by the relatively high score of realism, presence, and usability, as well as a low score of simulation sickness. Participant's responses regarding risk, safety, and trust in AV also confirmed their positive experience with AV in the virtual environment. Overall, the results of subjective measures suggest the feasibility of employing this type of VR system to investigate pedestrian-AV interaction in more complex traffic situations.

This study has several limitations that should be addressed in future research. Firstly, while the results of linear mixed model showed that random effects (i.e., participant's ID) had a significant impact on pedestrian road crossing behavior, we did not further investigate how personal demographics influence pedestrian-AV interaction in the current study. Moreover, although participant's trust in AV, PBC, and PR were collected in the questionnaire, it was not included in the mixed effect model. The future studies should explore the correlation between these factors and AV-pedestrian interaction to better understand how individual differences affect pedestrian road crossing behavior. Secondly, although we attempted to realistically resemble the representation of a shared space, more features should be considered and investigated in future studies, such as other modes of transportation. Thirdly, it is known that lighting and weather conditions can change the complexity of traffic environments. Future studies should include the factors examined in this study in different lighting and weather conditions. Moreover, the current study only considered pedestrian gazing behavior towards the vehicle, where future studies should further investigate pedestrian gazing behavior towards the eHMIs in order to more precisely understand the effect of visual attention in road crossing decisions.

6 CONCLUSIONS

To the author's knowledge, this is the first study that empirically investigated the interaction between multiple pedestrians and AVs across a wide range of conditions in shared spaces. We found that the presence of multiple pedestrians had a significant impact on pedestrian movement dynamics while crossing the road and their relative positions have a significant impact on the time before crossing and gazing behavior. Specifically, the pedestrians that were further away from the AV had less observation towards the AV and took longer time before making crossing decisions. In terms of the impact of AV's driving style, we found that more abrupt deceleration behavior exhibited by the AV reduces the decision time to cross the road. Moreover, our findings show that eHMIs in shared spaces can reduce pedestrian's visual attention toward the AV but not necessarily reduce the decision-making time to cross the road. By comparing pedestrian crossing between the straight path and the T-junction path, we found pedestrians were more cautious in more complex traffic scenarios. Furthermore, we found space gap plays an important role in pedestrian crossing decisions and movement dynamics during crossing. Our results show that this unique VR setup, namely multiple users, immersive setting, and real-walking locomotion offers a more comprehensive and objective approach to studying pedestrian-AV interaction in more complex traffic conditions.

ACKNOWLEDGMENTS

This work was supported by the project "Safe Interaction of Pedestrians and Cyclists with Automated Transport (SIPCAT)" founded by the Netherlands Organization for Scientific Research (Grant: 439.20.604). The authors thank The New Base for assisting in developing the VR scenarios. Moreover, we would like to thank AMS Institute and Marineterrein Amsterdam for their facilitation of the project.

REFERENCES

[1] J. Bindschädel, I. Krems, A. Kiesel, Interaction between pedestrians and automated vehicles: Exploring a motion-based approach for virtual reality experiments, Transp. Res. Part F Traffic Psychol. Behav. 82 (2021) 316–332. https://doi.org/10.1016/j.trf.2021.08.018.

[2] D. Dey, A. Van Vastenhoven, R.H. Cuijpers, M. Martens, B. Pfleging, Towards scalable eHMIs: Designing for AV-VRU communication beyond one pedestrian, in: Proc. - 13th Int. ACM Conf. Automot. User Interfaces Interact. Veh. Appl. AutomotiveUI 2021, 2021: pp. 274–286. https://doi.org/10.1145/3409118.3475129.

[3] L. Fridman, B. Mehler, L. Xia, Y. Yang, L.Y. Facusse, B. Reimer, To Walk or Not to Walk: Crowdsourced Assessment of External Vehicle-to-Pedestrian Displays, (2017). http://arxiv.org/abs/1707.02698.

[4] R. Masuda, S. Ono, T. Hiraoka, Y. Suda, Trust Estimation for Autonomous Vehicles by Measuring Pedestrian Behavior in VR, Companion 2023 ACM/IEEE Int. Conf. Human-Robot Interact. (2023) 203–207. https://doi.org/10.1145/3568294.3580072.

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

[6] A.C. Hensch, I. Neumann, M. Beggiato, J. Halama, J.F. Krems, How Should Automated Vehicles Communicate? - Effects of a Light-Based Communication Approach in a Wizard-of-Oz Study, Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-20503-4_8.

[7] E. Fratini, R. Welsh, P. Thomas, Ranking Crossing Scenario Complexity for eHMIs Testing: A Virtual Reality Study, Multimodal Technol. Interact. 7 (2023). https://doi.org/10.3390/mti7020016.

[8] N. Merat, T. Louw, R. Madigan, M. Wilbrink, A. Schieben, What externally presented information do VRUs require when interacting with fully Automated Road Transport Systems in shared space?, Accid. Anal. Prev. 118 (2018) 244–252. https://doi.org/10.1016/j.aap.2018.03.018.

[9] Y. Song, Q. Jiang, W. Chen, X. Zhuang, G. Ma, Pedestrians' road-crossing behavior towards eHMI-equipped autonomous vehicles driving in segregated and mixed traffic conditions, Accid. Anal. Prev. 188 (2023) 107115. https://doi.org/10.1016/j.aap.2023.107115.

[10] Y. Feng, D. Duives, W. Daamen, S. Hoogendoorn, Data collection methods for studying pedestrian behaviour: A systematic review, Build. Environ. 187 (2021) 107329. https://doi.org/https://doi.org/10.1016/j.buildenv.2020.107329.

[11] T.T.M. Tran, C. Parker, M. Tomitsch, A Review of Virtual Reality Studies on Autonomous Vehicle-Pedestrian Interaction, IEEE Trans. Human-Machine Syst. 51 (2021) 641–652. https://doi.org/10.1109/THMS.2021.3107517.

[12] S. Stadler, H. Cornet, T. Novaes Theoto, F. Frenkler, A Tool, not a Toy: Using Virtual Reality to Evaluate the Communication Between Autonomous Vehicles and Pedestrians, Augment. Real. Virtual Real. Power AR VR Bus. (2019) 203–216. https://doi.org/10.1007/978-3-030-06246-0_15.

[13] J.P.N. Velasco, Y. Mun Lee, J. Uttley, A. Solernou, H. Farah, B. van Arem, M. Hagenzieker, N. Merat, Will pedestrians cross the road before an automated vehicle? The effect of drivers' attentiveness and presence on pedestrians' road crossing behavior, Transp. Res. Interdiscip. Perspect. 12 (2021) 100466. https://doi.org/10.1016/j.trip.2021.100466.

[14] K. Tian, A. Tzigieras, C. Wei, Y.M. Lee, C. Holmes, M. Leonetti, N. Merat, R. Romano, G. Markkula, Deceleration parameters as implicit communication signals for pedestrians' crossing decisions and estimations of automated vehicle behaviour, Accid. Anal. Prev. 190 (2023) 107173. https://doi.org/10.1016/j.aap.2023.107173.

[15] I. Kaparias, M.G.H. Bell, A. Miri, C. Chan, B. Mount, Analysing the perceptions of pedestrians and drivers to shared space, Transp. Res. Part F Traffic Psychol. Behav. 15 (2012) 297–310. https://doi.org/10.1016/j.trf.2012.02.001.

[16] C.G. Burns, L. Oliveira, V. Hung, P. Thomas, S. Birrell, Pedestrian Attitudes to Shared-Space Interactions with Autonomous Vehicles - A Virtual Reality Study, Adv. Intell. Syst. Comput. 964 (2020) 307–316. https://doi.org/10.1007/978-3-030-20503-4_29.

[17] D. Dey, A. Matviienko, M. Berger, M. Martens, B. Pfleging, J. Terken, Communicating the intention of an automated vehicle to pedestrians: The contributions of eHMI and vehicle behavior, IT - Inf. Technol. 63 (2021) 123–141. https://doi.org/10.1515/itit-2020-0025.

[18] A.C. Hensch, I. Neumann, M. Beggiato, J. Halama, J.F. Krems, Effects of a light-based communication approach as an external HMI for Automated Vehicles - A Wizard-of-Oz Study, Trans. Transp. Sci. 10 (2019) 18–32. https://doi.org/10.5507/TOTS.2019.012.

[19] Y.E. Song, C. Lehsing, T. Fuest, K. Bengler, External HMIs and their effect on the interaction between pedestrians and automated vehicles, Adv. Intell. Syst. Comput. 722 (2018) 13–18. https://doi.org/10.1007/978-3-319-73888-8_3.

[20] Y. Feng, H. Farah, B. van Arem, Effect of eHMI on pedestrian road crossing behavior in shared space with Automated Vehicles - A Virtual Reality study, ArXiv Prepr. (2023). https://doi.org/arXiv:2308.05654.

[21] D. Dey, A. van Vastenhoven, R.H. Cuijpers, M. Martens, B. Pfleging, Towards Scalable eHMIs: Designing for AV-VRU Communication Beyond One Pedestrian, in: Proc. - 13th Int. ACM Conf. Automot. User Interfaces Interact. Veh. Appl. AutomotiveUI 2021, 2021: pp. 274–286. https://doi.org/10.1145/3409118.3475129.

[22] C. Ackermann, M. Beggiato, S. Schubert, J.F. Krems, An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles?, Appl. Ergon. 75 (2019) 272–282. https://doi.org/10.1016/j.apergo.2018.11.002.

[23] S. Deb, L.J. Strawderman, D.W. Carruth, Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment, Transp. Res. Part F Traffic Psychol. Behav. 59 (2018) 135–149. https://doi.org/10.1016/j.trf.2018.08.016.

[24] K. de Clercq, A. Dietrich, J.P. Núñez Velasco, J. de Winter, R. Happee, External Human-Machine Interfaces on Automated Vehicles: Effects on Pedestrian Crossing Decisions, Hum. Factors. 61 (2019) 1353–1370. https://doi.org/10.1177/0018720819836343.

[25] A. Habibovic, V.M. Lundgren, J. Andersson, M. Klingegård, T. Lagström, A. Sirkka, J. Fagerlönn, C. Edgren, R. Fredriksson, S. Krupenia, D. Saluäär, P. Larsson, Communicating intent of automated vehicles to pedestrians, Front. Psychol. 9 (2018). https://doi.org/10.3389/fpsyg.2018.01336.

[26] S.M. Faas, L.A. Mathis, M. Baumann, External HMI for self-driving vehicles: Which information shall be displayed?, Transp. Res. Part F Traffic Psychol. Behav. 68 (2020) 171–186. https://doi.org/10.1016/j.trf.2019.12.009.

[27] I. Othersen, Antonia S. Conti-Kufner, A. Dietrich, P. Maruhn, K. Bengler, Designing for Automated Vehicle and Pedestrian Communication: Perspectives on eHMIs from Older and Younger Persons, in: Proc. Hum. Factors Ergon. Soc. Eur. Chapter 2018 Annu. Conf., 2018: pp. 135–148.

[28] T.T. Nguyen, K. Holländer, M. Hoggenmueller, C. Parker, M. Tomitsch, Designing for Projection-based Communication between Autonomous Vehicles and Pedestrians, in: Proc. - 11th Int. ACM Conf. Automot. User Interfaces Interact. Veh. Appl. AutomotiveUI 2019, 2019: pp. 284–294. https://doi.org/10.1145/3342197.3344543.

[29] A. Löcken, C. Golling, A. Riener, How should automated vehicles interact with pedestrians? A comparative analysis of interaction concepts in virtual reality, Proc. - 11th Int. ACM Conf. Automot. User Interfaces Interact. Veh. Appl. AutomotiveUI 2019. (2019) 262–274. https://doi.org/10.1145/3342197.3344544.

[30] A.H. Kalantari, Y. Yang, J.G. De Pedro, Y.M. Lee, A. Solernou, C. Holmes, N. Merat, G. Markkula, Who goes first ? A distributed simulator study of vehicle-pedestrian interaction, Accid. Anal. Prev. 186 (2023) 107050. https://doi.org/10.1016/j.aap.2023.107050.

[31] Y. Li, M. Dikmen, T.G. Hussein, Y. Wang, C. Burns, To cross or not to cross: Urgency-based external warning displays on autonomous vehicles to improve pedestrian crossing safety, in: Proc. - 10th Int. ACM Conf. Automot. User Interfaces Interact. Veh. Appl. AutomotiveUI 2018, 2018: pp. 188–197. https://doi.org/10.1145/3239060.3239082.

[32] K. Mahadevan, E. Sanoubari, S. Somanath, J.E. Young, E. Sharlin, AV-pedestrian interaction design using a pedestrian mixed traffic simulator, DIS 2019 - Proc. 2019 ACM Des. Interact. Syst. Conf. (2019) 475–486. https://doi.org/10.1145/3322276.3322328.

[33] A.N. Rakotoarivelo, J.M. Auberlet, R. Brémond, Will I cross when I see you crossing? Towards illegal pedestrian road crossing simulations, Transp. Res. Part F Traffic Psychol. Behav. 93 (2023) 309–325. https://doi.org/10.1016/j.trf.2023.01.022.

[34] S. Das, C.F. Manski, M.D. Manuszak, Walk or wait? An empirical analysis of street crossing decisions, J. Appl. Econom. 20 (2005) 529–548. https://doi.org/10.1002/jae.791.

[35] Y. Jiang, E.E. O'Neal, S. Zhou, J.M. Plumert, J.K. Kearney, Crossing roads with a computer-generated agent: Persistent effects on perception-action tuning, ACM Trans. Appl. Percept. 18 (2021) 1–16. https://doi.org/10.1145/3431923.

[36] A. Koilias, M.G. Nelson, C.N. Anagnostopoulos, C. Mousas, Immersive walking in a virtual crowd: The effects of the density, speed, and direction of a virtual crowd on human movement behavior, Comput. Animat. Virtual Worlds. 31 (2020) 1–14. https://doi.org/10.1002/cav.1928.

[37] Y. Jiang, P. Rahimian, E.E. O'Neal, J.M. Plumert, J.P. Yon, J.K. Kearney, L. Franzen, Acting together: Joint pedestrian road crossing in an immersive virtual environment, ACM Trans. Appl. Perception(TAP). 15 (2016) 1–13. https://doi.org/10.1109/VR.2016.7504719.

[38] E.E. O'Neal, Y. Jiang, K. Brown, J.K. Kearney, J.M. Plumert, How Does Crossing Roads with Friends Impact Risk Taking in Young Adolescents and Adults?, J. Pediatr. Psychol. 44 (2019) 726–735. https://doi.org/10.1093/jpepsy/jsz020.

[39] M. Hoggenmueller, M. Tomitsch, S. Worrall, Designing Interactions With Shared AVs in Complex Urban Mobility Scenarios, Front. Comput. Sci. 4 (2022) 1–14. https://doi.org/10.3389/fcomp.2022.866258.

[40] W. Tabone, Y.M. Lee, N. Merat, R. Happee, J. de Winter, Towards future pedestrian-vehicle interactions: Introducing theoretically-supported AR prototypes, in: 13th Int. Conf. Automot. User Interfaces Interact. Veh. Appl., 2021: pp. 209–218. https://doi.org/10.1145/3409118.3475149.

[41] H. Furuya, K. Kim, G. Bruder, P.J. Wisniewski, G.F. Welch, Autonomous Vehicle Visual Embodiment for Pedestrian Interactions in Crossing Scenarios: Virtual Drivers in AVs for Pedestrian Crossing, in: Conf. Hum. Factors Comput. Syst. - Proc., 2021: pp. 1–7. https://doi.org/10.1145/3411763.3451626.

[42] B.E. Riecke, B. Bodenheimer, T.P. McNamara, B. Williams, P. Peng, D. Feuereissen, Do we need to walk for effective virtual reality navigation? Physical rotations alone may suffice, in: Int. Conf. Spat. Cogn., 2010: pp. 234–247. https://doi.org/10.1007/978-3-642-14749-4_21.

[43] P.T. Wilson, W. Kalescky, A. MacLaughlin, B. Williams, VR locomotion: Walking>Walking in Place>Arm swinging, in: Proc. 15th ACM SIGGRAPH Conf. Virtual-Reality Contin. Its Appl. Ind., 2016: pp. 243–249. https://doi.org/10.1145/3013971.3014010.

[44] M.S. Martin Usoh, A. Steed, Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality, ACM Trans. Comput. Interact. 2 (1995) 201–219. https://doi.org/10.1145/210079.210084.

[45] N.A. Kaptein, J. Theeuwes, R. van der Horst, Driving simulator validity: some considerations, Transp. Res. Rec. (1996) 30–36. https://doi.org/10.3141/1550-05.
[46] R.S. Kennedy, N.E. Lane, K.S. Berbaum, M.G. Lilienthal, Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness, Int. J. Aviat. Psychol. 3 (1993) 203–220. https://doi.org/10.1207/s15327108ijap0303_3.
[47] B.G. Witmer, C.J. Jerome, M.J. Singer, The Factor Structure of the Presence Questionnaire, Presence Teleoperators Virtual Environ. 14 (2005) 298–312.
[48] J.P. Nuñez Velasco, H. Farah, B. van Arem, M.P. Hagenzieker, Studying pedestrians' crossing behavior when interacting with automated vehicles using virtual reality, Transp. Res. Part F Traffic Psychol. Behav. 66 (2019) 1–14. https://doi.org/10.1016/j.trf.2019.08.015.
[49] J. Brooke, SUS - A quick and dirty usability scale, Usability Eval. Ind. 189 (1996) 4–7. https://doi.org/10.1002/hbm.20701.
[50] D.C. Schwebel, J. Gaines, J. Severson, Validation of virtual reality as a tool to understand and prevent child pedestrian injury, Accid. Anal. Prev. 40 (2008) 1394–1400. https://doi.org/10.1016/j.aap.2008.03.005.
[51] W. Payre, J. Cestac, P. Delhomme, Fully Automated Driving: Impact of Trust and Practice on Manual Control Recovery, Hum. Factors. 58 (2016) 229–241. https://doi.org/10.1177/0018720815612319.
[52] J.P. Nuñez Velasco, H. Farah, B. van Arem, M.P. Hagenzieker, J.P.N. Velasco, H. Farah, B. Van Arem, M.P. Hagenzieker, Studying pedestrians' crossing behavior when interacting with automated vehicles using virtual reality, Transp. Res. Part F Traffic Psychol. Behav. 66 (2019) 1–14. https://doi.org/10.1016/j.trf.2019.08.015.
[53] A. Bangor, P. Kortum, J. Miller, Determining What Individual SUS Scores Mean : Adding an Adjective Rating Scale, J. Usability Stud. 4 (2009) 114–123.
[54] J.P.N. Velasco, Y. Mun Lee, J. Uttley, A. Solernou, H. Farah, B. van Arem, M. Hagenzieker, N. Merat, Will pedestrians cross the road before an automated vehicle? The effect of drivers' attentiveness and presence on pedestrians' road crossing behavior, Transp. Res. Interdiscip. Perspect. 12 (2021) 100466. https://doi.org/10.1016/j.trip.2021.100466.