



Investigating pedestrian crossing decision with autonomous cars in virtual reality

Shuaixin Qi¹ · Marino Menozzi¹

Accepted: 13 March 2023

© The Author(s) 2023

Abstract

With the development of autonomous vehicle (AV) technology, understanding how pedestrians interact with AVs is of increasing importance. In most field studies on pedestrian crossing behavior when encountering AVs, pedestrians were not permitted to physically cross the street due to safety restrictions. Instead, the physical crossing experience was replaced with indirect methods (e.g., by signaling with gestures). We hypothesized that this lack of a physical crossing experience could influence the participants' crossing behavior. To test this hypothesis, we adapted a reference study and constructed a crossing facility using a virtual reality (VR) simulation. In a controlled experiment, the participants encountered iterations of oncoming AVs. For each interaction, they were asked to either cross the street or signify their crossing decisions by taking steps at the edge of the street without crossing.

Our study reveals that the lack of a physical crossing can lead to a significantly lower measured critical gap and perceived stress levels, thus indicating the need for detailed analysis when indirect methods are applied for future field studies.

Practical Relevance: Due to safety requirements, experiments will continue to measure participants' crossing behavior without permitting them to physically walk in front of an oncoming vehicle. Our study was the first attempt to reveal how this lack of crossing could potentially affect pedestrians' behavior, and we obtained empirical evidence in support of our hypothesis, thus providing insights for future studies.

Keywords Virtual Reality · Autonomous vehicle · Pedestrian behavior · Gap acceptance · Road safety

Untersuchung des Entscheidungsverhaltens von Fußgängern bei Überqueren mit autonomen Fahrzeugen in virtueller Realität

Zusammenfassung

Mit der Entwicklung der autonomen Fahrzeuge (AVs) gewinnt das Verständnis für die Interaktion von AVs mit Fußgängern zunehmend an Bedeutung. In den meisten Feldstudien zum Überquerungsverhalten von Fußgängern bei der Begegnung mit AVs war es aufgrund von Sicherheitsbeschränkungen den Fußgängern nicht gestattet, die Straße physisch zu überqueren. Stattdessen wurden die Erfahrungen von physischen Überquerungen durch indirekte Methoden ersetzt, beispielsweise durch Signalisierung mit Gesten. Es wird der Hypothese nachgegangen, dass dieser Mangel an physischer Überquerungserfahrung das Entscheidungsverhalten der Teilnehmer beeinflussen könnte. Eine Referenzstudie wurde adaptiert und eine Überquerungsanlage mit Hilfe einer Virtual-Reality-Simulation (VR) konstruiert. In einem kontrollierten Experiment wurden die Teilnehmer mit AVs konfrontiert. Bei jeder Interaktion wurden sie gebeten, entweder die Straße physisch zu überqueren oder ihre Überquerungsentscheidung durch Schritte am Straßenrand zu signalisieren.

Unsere Untersuchung zeigt, dass ein Fehlen der Überquerungserfahrung zu einem signifikant geringeren kritischen Abstand und zu höheren empfundenen Stressniveaus führen kann. Dies verdeutlicht die Notwendigkeit einer detaillierten Analyse bei der Anwendung indirekter Methoden in zukünftigen Feldstudien.

✉ Masch.-Ing. Shuaixin Qi, BSc
qis@ethz.ch

¹ Eidgenössische Technische Hochschule Zürich (ETHZ),
Scheuchzerstrasse 7, 8092 Zürich, Switzerland

Praktische Relevanz: Zum ersten Mal konnte nachgewiesen werden, dass das Entscheidungsverhalten von Fußgängern durch physisches Überqueren in VR sich von jenem unterscheidet, welches durch Anzeigen der Bereitschaft, die Straße zu überqueren erfasst wird. Diese Erkenntnis ist für künftige Untersuchungen zu berücksichtigen.

Schlüsselwörter Virtuelle Realität · Autonomes Fahren · Fußgängerverhalten · Verkehrssicherheit

1 Introduction

Studies on autonomous vehicle (AV) technology represent a growing field and have attracted the attention of both researchers and policymakers. It is considered that AVs have the potential to increase traffic safety and sustainability (Fagnant and Kockelman 2015; Van Brummelen et al. 2018). However, AVs will likely share traffic space with conventional traffic users, including pedestrians. Established communication cues between pedestrians and drivers, such as eye contact (Sucha et al. 2017) or body movements (Schmidt and Färber 2009), may become less reliable as AV drivers may be distracted or even absent (Mahadevan et al. 2018). Therefore, it is vital to understand pedestrian crossing behavior in response to oncoming AVs.

Though considerable literature has been published on pedestrian crossing behavior, interactions between pedestrians and AVs are not yet fully understood (Habibovic et al. 2016). Since highly automated vehicles (SAE level 4/5) (On-Road Automated Driving (ORAD) committee 2016) are not yet available, alternative methods are required to observe pedestrian behavior and collect data. In this regard, various novel methods have been explored with a focus on methods for full-scale field experiments.

Firstly, we reviewed the general approaches used in studies on pedestrian-AV interaction. One commonly used approach to studying interactions between pedestrians and AVs is the Wizard of Oz (WoZ) technique. In this method, participants are asked to interact with an AV while researchers operate the vehicle or its interfaces while hidden from sight. One of the first realizations of this technique was proposed by Rothenbücher et al. (2016), where a human driver was concealed behind a car seat costume (i.e., “the ghost driver”) to simulate a driverless vehicle. Their observations on real roads showed that most pedestrians could seamlessly cross the street without any explicit external human-machine interfaces (eHMI). Other studies have since used the ghost driver method and found that most pedestrians may not require dedicated signaling cues apart from the motion of the AV (Currano et al. 2018; Moore et al. 2019). In addition, researchers have also used a programmable vehicle with a scale of roughly 1:12 of a Mercedes-Benz Smart car as an oncoming vehicle to evoke emotional responses from pedestrians (Zimmermann and Wettach 2017). While this approach may not be as prac-

tical for large-scale field experiments, it can be useful for studying specific aspects of pedestrian behavior.

Quantitatively evaluating real-life interactions between a pedestrian and an AV is often challenging in a controlled experiment with real-world traffic conditions. Due to safety and ethical requirements, prior field studies have rarely allowed participants to physically cross the road in front of an operating vehicle. Instead, researchers have come up with alternative approaches to assess pedestrian interaction.

We explored some of the most recent practices and how the lack of physical crossing was addressed: To start with, Joisten et al. (2020) measured participants’ crossing behavior in terms of critical gap acceptance and perceived safety with a simulated WoZ AV. However, the participants were not permitted to step in front of the vehicle for safety reasons. Walker et al. (2019) developed a slider as an input device to assess participants’ willingness to cross in real-time in response to an oncoming AV. Dey et al. (2021) utilized this approach to study pedestrian-AV interaction and found that eHMI could increase pedestrians’ willingness to cross in low-speed situations. However, the possibility that participants exhibited greater risk-taking behavior due to the lack of a physical crossing was named as one of the study’s limitations. Mahadevan et al. (2018) studied the effect of eHMI prototypes on a simulated WoZ car and a Segway, where the participants were permitted to cross in front of the Segway but were only permitted to express their crossing intention indirectly when interacting with the oncoming WoZ car. Rodríguez Palmeiro et al. (2018) measured participants’ crossing behavior in response to both traditional and automated WoZ vehicles by asking them to step forwards at the beginning of the experiment until the last moment they deemed acceptable to cross when they were to step backward. The results revealed no significant differences in the critical gap acceptance or self-reported stress between the different vehicle conditions, though participants subjectively reported having been influenced by this. The lack of a physical crossing was also suspected of having influenced the overall result. From these existing studies, however, the ways in which the lack of crossing might have influenced pedestrian crossing behavior are not evident. Furthermore, this issue may persist for the foreseeable future until studies can safely involve participants walking in front of an oncoming vehicle within the parameters of safety and ethical limitations.

These same limitations do not apply to virtual reality (VR)-based experiments, which are frequently used as an alternative to real-world setups to assess participants' behavior. To study the feasibility of VR technology, we also explored the characteristics of VR as a simulation method and its recent application in pedestrian studies. VR can recreate similar traffic interactions in a more controlled environment with no risk of traffic accidents (Bhagavathula et al. 2018). Among photographs and panorama, VR was shown to be the most realistic display format for measuring physiological responses (Higuera-Trujillo et al. 2017). Singh et al. (2015) confirmed that VR-based methodology could produce results that accurately predict certain perceptions of pedestrians to real-world vehicles. Deb et al. (2017) built a VR simulator to obtain objective pedestrian behavior data which matched real-world norms. Similarly, Bhagavathula et al. (2018) compared data from pedestrian-vehicle interactions in equivalent real and virtual environments, and the results suggested no significant differences. Although VR cannot perfectly replicate real-world scenarios, it has been proven to be a reasonable tool with which to identify significant factors affecting pedestrians' behavior and in observing general trends that are transferable to reality. As a matter of fact, some existing pedestrian-AV VR simulators already include physically crossing in front of an oncoming AV as part of the interaction (Deb et al. 2018; Kooijman et al. 2019; Löcken et al. 2019). However, to the best of our knowledge, there is a notable lack of studies that seek to identify the ways in which the lack of physical crossing influences pedestrian crossing behavior.

To summarize, the substantial advances in AV technology herald a greater need to understand how vulnerable street users such as pedestrians will interact with AVs. However, due to safety regulations, it is not currently possible to conduct field experiments on pedestrian crossing behavior using a physical crossing. Given the proven ability of VR to recreate realistic traffic situations, this study aims to bridge the knowledge gap with VR simulation.

Our hypothesis posits that the decision-making behavior of pedestrians varies depending on the manner in which they indicate their intention to cross, be it directly or indirectly. This variance in behavior will be quantified by objectively measuring critical gap acceptance and subjectively assessing perceived stress levels. Specifically, we expect that the physical crossing would lead to increased stress levels and higher critical gaps. To frame the scope of the study, we decided to orient it around the aforementioned experiment of Rodríguez Palmeiro et al. (2018), whose approach (i.e., stepping) can be replicated well in VR without extra input devices. We designed a controlled VR experiment where participants interacted with an oncoming AV. Their measured critical gap was calculated based on the recorded trajectories, and their perceived stress was evaluated after

each interaction. Further quantitative and qualitative data were also collected to determine how closely our measurements match real-world norms. Inspired by the WoZ aspect of prior field studies, we developed simple techniques to achieve adequate immersion and evaluated the effectiveness through retrospective questions. The experiment was approved by the ETH Ethics Commission (2021-N-213).

2 Methods

2.1 Participants

Twenty-four participants between 19 and 30 ($M=23.8$, $SD=2.4$) were recruited through social media and pinboards on the ETH Zurich university campus. Normal health conditions, a good command of English, and familiarity with right-hand traffic were required for the participation.

Interested individuals could register through a web form. As a part of the registration, candidates were screened to exclude those likely to suffer from motion sickness. For this purpose, the short version of the motion sickness susceptibility questionnaire (MSSQ-short) (Golding 1998) was used. Based on an estimation of the time spent in the simulation (8 min), a threshold score for the MSSQ-short of 34.1 was computed. This MSSQ score was calculated based on Fig. 3 in Golding (1998), which generated a score for the long version of the MSSQ. Further details on the computation method for the threshold and the conversion between MSSQ-long and MSSQ-short are given in the Procedure section of a previous study (Ropelato et al. 2022). Candidates with scores higher than the threshold were excluded from participation.

2.2 Instrumentation

A virtual intersection was implemented in Unity (2020.3.19f) and presented using a cabled HTC Vive Pro head-mounted display (HMD). Participants were able to move in the virtual world through naturalistic walking in the real world. The position and orientation of the HMD were tracked in real-time and recorded together with the position of the vehicle at a rate of 5 Hz.

Figure 1 shows an excerpt of the intersection as seen from the participants' perspective, and Fig. 2 is a bird's-eye view of the whole scene. A virtual environment (VE) built with high-resolution assets (Equilibrium A 2022; Black Starling Productions 2021; 255 Pixel Studios 2020) from the Unity asset store was used to ensure graphical fidelity. A green bus (Edy 2020) was placed on the scene as the destination of the crossing scenario. To prevent participants from bumping into the real-world wall while traversing the

Fig. 1 The side view of the crossing facility with a length of 7 m

Abb. 1 Seitenansicht des Fußgängerüberwegs mit einer Länge von 7 m



Fig. 2 The Bird's-eye view of the crossing facility. The vertical and horizontal arrows denote the paths of the pedestrian and the car, respectively

Abb. 2 Vogelperspektive auf dem Fußgängerüberweg. Die vertikalen und horizontalen Pfeile bezeichnen die Wege des Fußgängers bzw. des Autos

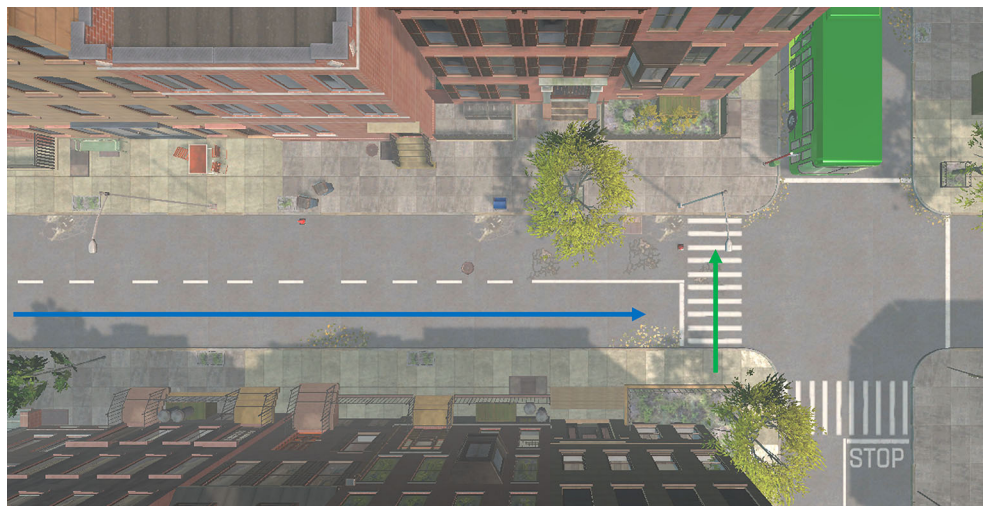


Fig. 3 *Left:* The virtual vehicle model with a roof sign; *Right:* The clips showing an attentive driver (*bottom right*) and an inattentive driver (*upper right*)

Abb. 3 *Links:* Das virtuelle Fahrzeugmodell mit einem „Self-driving“ Schild auf dem Dach; *Rechts:* Videoabschnitte eines aufmerksamen Fahrers (*unten*) und eines unaufmerksamen Fahrers (*oben*)



street in the VE, a virtual traffic cone was placed on the ground in the real world as a guide (shown in Fig. 1). Once participants reached the cone, a warning message appeared in the VE to prevent them from walking into the wall. Although participants did not walk all the way onto the bus, the most relevant aspects of the experience of crossing the street were faithfully conveyed.

In reality, a two-way road has a width of approximately 7 m (U.S. Department of Transportation, Federal Highway Administration 2007; Highways England 2016). Accordingly, we modeled a street of 7 m width in the VE. However, as the length of our facility was only 4.1 m, the translational gain of the HMD was doubled, i.e., the rate at which movements in the real world would translate to the participants' perspective in the virtual world was doubled. This technique has been shown to work well for creating larger VEs within smaller physical rooms (Ropelato et al. 2022; Williams et al. 2006). Since increasing the translational gain on all axes also amplifies slight bobbing and sideward tilting of natural head movements, only the gain parallel to the crossing facility (i.e., along the vertical arrow in Fig. 2) was increased.

In the foreseeable future, pedestrians may encounter various traffic scenarios where vehicles may be driven by a human driver or operate autonomously. Moreover, the attentiveness of the driver may vary, and the vehicle's behavior at pedestrian crossings may differ, with the option to either pass the crossing or yield to pedestrians. To accommodate these variables, we have developed two test scenarios: one with an attentive driver and another with an inattentive driver. In both scenarios, the vehicle may either stop or yield at the pedestrian crossing.

A medium-sized silver coupe (Pro 3D Models 2018) was used as the vehicle model for all trials. A roof sign with the message "self-driving" was attached to the vehicle (Fig. 3). The vehicle always approached from the left of the participants, as the approaching direction was shown to be insignificant to the crossing behavior (Rodríguez Palmeiro et al. 2018).

To increase the immersiveness, a 2D plane displaying a pre-recorded video clip of a human driver was placed above the driver's seat (Fig. 3). This was a low-complexity replacement for a real-time projection, as we wanted to keep the setup simple and avoid introducing additional variables.

We modified a vehicle controller (Edy's Vehicle Physics 2020) from the Unity assets store to execute the vehicle's trajectory. Two different trajectories were manually predefined according to the independent variable 'stopping behavior.' The acceleration profile was slightly varied for each interaction to make the resulting trajectories less robotic over the course of repeated interactions. Manual control of the vehicle was intentionally avoided, as this would increase the complexity of the setup and introduce bias.

Following the results of a pilot study, the possibility of collision as a consequence of the interaction was eliminated, as six out of the seven volunteers in the pilot study indicated that collisions were scary. Furthermore, when the collision outcome was tested, four volunteers attempted an abrupt dodging maneuver to avoid the collision in VE and risked bumping into the wall or the experimenter. To avoid collisions, an invisible collider box of 5 m was attached to the front of the vehicle, and the vehicle model was disabled and therefore disappeared as soon as the collider box touched the participant.

To prevent participants from memorizing the vehicle's trajectory, the model of the vehicle was disabled once the participant had passed the midline of the crossing path.

2.3 Independent variables

The main aim of our study was to investigate the potential effect of the lack of physical crossing in VR pedestrian-AV interactions on pedestrians' behavior. Two crossing methods (CMs) were thus considered: The first CM replicated the method used in a reference field study (Rodríguez Palmeiro et al. 2018), in which the participants were instructed to take one step forward at the first moment they would cross the road and one step backward at the last moment they would cross the road, as in the aforementioned study. This method will hereafter be referred to as the 'stepping method.' For the second CM, participants were asked to cross the road at the last moment they would cross. This method will hereafter be referred to as the 'crossing method.'

We expected to measure similar behavior if the CM were not a significant factor by enforcing the same last-moment constraint for both CMs. These constraints were also needed for the implicit measurement of one of the dependent variables (the critical gap), which will be elaborated on later. Similar to the reference study, we included the driver's attentional state and the stopping behavior of the vehicle as moderating variables in addition to the CM.

The effect of the AV driver's attentivity on the interaction was evaluated by using video clips of two different drivers: one attentive and one inattentive. As explained above, the clips of the driver presented in the VE were pre-recorded. For the attentive cases, the video clip shows a driver who holds the steering wheel and looks intently at the traffic in front of him. On approaching the crossing, the driver looks into the camera to seek eye contact with the pedestrian. With the inattentive driver, the driver operates a smartphone with both hands. His gaze is intentionally kept away from the camera to signify that he is not paying attention to traffic.

Two different stopping behaviors of the car were investigated, in which the speed profile and the stopping position in front of the pedestrian crossing were varied. This resulted

in a yielding outcome and a non-yielding outcome. For the yielding case, the vehicle started 60 m from the crossing and accelerated to a speed of about 33–37 km/h before gradually coming to a full stop in front of the pedestrian crossing. For the non-yielding case, the vehicle started in the same position and accelerated to a speed of about 31–34 km/h, after which it only braked slightly before the crossing and then passed the crossing without yielding.

Taking into account all previous within-subject variables, we have eight possible combinations ($2 \times 2 \times 2$) at this stage, and each participant should experience each of these interactions twice (giving 16 iterations total). However, frequently switching between different CMs may confuse participants. Therefore, we decided to let participants start with either CM and go through the corresponding four combinations. Then, they were switched to the second CM to complete the first half (i.e., the first block) of the experiment. Afterward, starting with the second CM, the participant went through the second block of the experiment, following the same procedure as the first block. A fourth variable called “block number” was introduced to note to which half of the experiment each interaction belonged, and we expect to reveal possible learning effects through this.

Lastly, a between-subject variable called “trial number” was defined, denoting the method with which participants started to examine possible sequence effects.

2.4 Dependent variables

Two dependent variables were included in the study to evaluate the participants’ crossing behavior: the critical gap and the self-reported stress level. These two variables were also included in the reference study, and the findings will be compared. The critical gap is defined as the shortest time in seconds that can be accepted for pedestrians to cross the existing traffic condition (Brewer et al. 2006; National Research Council 2010). In our study, the critical gap is measured in terms of the distance between the bumper of the vehicle and the pedestrian at the moment a decision is taken to cross the street. In the stepping method, this is the moment the participant takes a step backward, and in the crossing method, this is the moment when the participant starts to cross the street. The critical gap was evaluated manually by inspecting the positions at these moments of the pedestrian and the vehicle, which were recorded at a rate of 5 Hz.

Rather than utilizing a physiological approach, self-assessments were employed to capture stress levels following each interaction, measured on a Likert scale ranging from 0–10, where 0 represented a state of relaxation and 10 indicated a highly stressed state. This method was adopted to achieve results that were comparable to the reference study.

We conducted interviews with partly identical questions to the reference study (Rodríguez Palmeiro et al. 2018) to obtain comparable results. After each completed interaction, a short oral interview (post-interaction) was conducted. After all interactions had been completed, a final oral interview (post-experiment) was carried out, followed by a final written questionnaire. The experimenter documented the answers to the oral interviews in written form, and the answers to the written questionnaires were collected from the participants using Google forms. The participants were permitted to refer to their previous answers for repeated questions in the event their answers were the same.

The post-interaction interview included two of the three questions from the reference study:

1. Which factors did you take into account before deciding to take a step backward/to cross the street (multiple answers allowed)?
2. How stressed were you on a scale from 0, not stressed at all, to 10, extremely stressed?

The post-experiment interview consisted of one question (no. 1) from the reference study and two purpose-phrased questions for evaluating participants’ feelings towards both CM:

1. How realistic do you think the setup of this experiment was on a scale from 0 (not realistic at all) to 10 (highly realistic)? Do you think it was similar to a crossing situation in real life? (Why/why not?)
2. On a scale from 0 (unnatural) to 10 (very natural), how natural was it for you to take a step forward at the first moment you would cross the road and a step backward at the last moment you would cross? (What are the reasons for your rating?)
3. On a scale from 0 (unnatural) to 10 (very natural), how natural was it for you to cross the street at the last moment? (What are the reasons for your rating?)

The final written questionnaire consisted of two questions to assess how well our implementation of a low-complexity AV worked. Lastly, we used a modified version of the four-factor model from the established presence questionnaire (PQ) (Witmer et al. 2005) to measure participants’ perceived presence in our VE, as well as their general ability to become immersed in the simulation.

1. The pre-recorded driver: Did you realize that the projection of the driver’s face was actually pre-recorded? In case of an affirmative answer: Did this influence your decision-making, and if yes, how?
2. The AV’s trajectory: Did you realize that the trajectory of the vehicle was not controlled by a human driver but was pre-programmed? In case of an affirmative answer: Did this influence your decision-making, and if yes, how?

3. PQ with questions 1–8, 11, 12, 14, 19, 21–25, 30, and 31 from Table 1 of Witmer et al. (2005).

2.5 Experiment procedure

Once interested individuals had completed the registration form, we evaluated their MSSQ-short score and arranged appointments for those who passed. Once participants had arrived at the experiment facility, we explained the steps of the experiment and collected signed consent forms. The participants were then led to the marked starting position, and the headset was fitted with the experimenter's assistance.

In the VE, participants started on the street edge, facing the non-signalized crossing facility. The vertical arrow in Fig. 2 visualizes their starting position and path. The vehicle approached from the participant's left and traversed the path shown by the horizontal arrow in Fig. 2. The experimenter explained both CMs to the participants. The participants were informed that the vehicle was sometimes actively controlled by the driver and that the vehicle would not always yield in order to increase the chance that their reactions would be more realistic. They were then given time to move around in the VE and experience both CMs in a trial scene until they felt confident to proceed. The trial scene comprised the same environment as the scenes in the experiment, with one exception: a different vehicle type, i.e., a Jeep with dark windows, was used, and thus no drivers were visible in the trial scene. This was because we wanted the participants to focus on learning the environment and how they could control their perspective rather than other features, such as the driver's attentional state.

Once the trial was completed, the experiment started. The experimenter remained in proximity to the participants at all times to be able to intervene to prevent tripping or collision with walls. At the end of the experiment, the experimenter provided a debriefing to explain the withdrawn information about the pre-recorded driver and the programmed vehicle trajectory.

During the experiment, nobody felt sick or aborted the experiment. No injuries or accidents occurred during the entire study. No collisions were observed in the VR.

3 Results

3.1 Critical gap and stress level

A total of 20 complete sets of critical gap data were collected from a sample population. In addition, we obtained 16 complete sets of self-reported stress levels from a subset of this sample. The remaining incomplete sets of data were excluded from the analysis, as they were missing one

measurement due to data loss or incorrectly executed experiment sequences. The exclusion was necessary to ensure the credibility and accuracy of the data. Interpolation was not performed on the incomplete sets to avoid potential biases and errors in the final analysis. Using the software SPSS, we performed two repeated measures analysis of variance (ANOVA) for each dependent variable. The analysis revealed that only the CM factor had a significant effect on the dependent variables, while the driver's attentivity, stopping behavior of the vehicle, and block number did not demonstrate a statistically significant impact. Since there were only two levels for the CM, no Mauchly's sphericity test was performed, and the Greenhouse-Geisser corrected tests are reported:

- The critical gap was significantly affected by the CM with $F(1)=45.383$, $p<0.001$, partial $\eta^2=0.706$.
- Similarly, the stress level was also significantly affected by the CM with $F(1)=56.945$, $p<0.001$, partial $\eta^2=0.803$.

From the estimated marginal means, it can be speculated that participants experienced higher stress and required a larger critical gap when they actually crossed the road (Fig. 4).

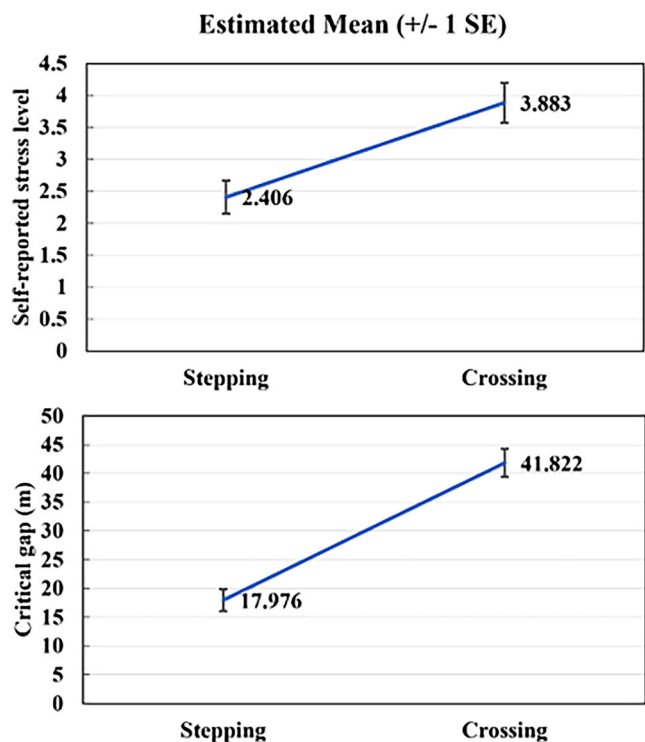


Fig. 4 Estimated mean for both dependent variables

Abb. 4 Die geschätzten Randmittelwerte für beide abhängigen Variablen

Table 1 PQ results categorized by subscale ($N=18$)**Tab. 1** Kategorisierte Ergebnisse von PQ

	Mean	Mean in percent	Median	Interquartile range
<i>Involvement (8–56)</i>	42.1	75	42.5	5
<i>Sensory fidelity (3–21)</i>	16.4	78	17	5
<i>Adaptivity/Immersion (5–35)</i>	28.8	82.3	29.5	6
<i>Interface quality (3–21)</i>	17	81	17	2

3.2 Factors for crossing decisions

The ratings from two participants were incomplete due to data loss, and therefore, only the results from the remaining 18 participants were considered for Sect. 3.3, 3.4, 3.5, and 3.6. In response to the question, “Which factors did you take into account before making the decision to take a step backward/to cross the street?” three hundred twenty answers were received (i.e., 20 participants \times 16 interactions), reporting 770 factors in total (multiple answers allowed). The most mentioned factors were the speed of the vehicle (230 out of 770) and the distance to the vehicle (169 out of 770). Other frequently reported factors in descending order included: the yielding behavior of the vehicle (80 out of 770), the sound of the vehicle (59 out of 770), the driver’s attentivity (21 out of 770), and eye contact or lack thereof, with the driver (20 out of 770).

3.3 Level of realism

On a scale of 0–10, our setup received a mean score of 6.9 ($SD=2.1$, $N=18$) for the level of realism. Only three participants answered “No” to the question, “Do you think it was similar to a crossing situation in real life?” The first person claimed that the driver’s face was not visible until much closer than in real life. The second found it difficult to relate the experiment to reality as he/she had never encountered a real AV. The third and final participant complained that they found the last-moment constraint counterintuitive. By contrast, seven out of 15 participants who gave positive responses complimented the combination of good visual and audio cues, while three found the vehicle’s dynamic to be realistic.

3.4 Naturalness of crossing methods

Both CMs received suboptimal ratings on their naturalness. The stepping method had a slightly higher average score ($M=6.5$, $SD=3.0$, $N=18$) than the crossing method ($M=5.3$, $SD=3.1$, $N=18$), and a moderate Pearson correlation ($r=0.4$) can be observed. A paired t-test yields a two-tailed p -value of 0.162; thus, the difference is not statistically significant. The reason most frequently given for the low rating across both methods was the fact that participants found it difficult to determine the last acceptable crossing moment for themselves. Three participants claimed they felt

more engaged with the stepping method because they could express their intentions through their movements. Two participants gave high ratings for the stepping method and low ratings for the crossing method because they perceived a greater risk during the crossing interactions.

3.5 Manipulation check

Thirteen participants noticed that the videos of the drivers were pre-recorded, but only three claimed that the realization influenced their decision-making. One person claimed to have paid little attention to the driver before crossing, and another noted that they stopped to make decisions based on the driver’s video. By contrast, only nine participants realized that the vehicles’ trajectory was not controlled by a human driver, and none claimed to be affected by this realization when making their crossing decision.

3.6 Presence questionnaire (PQ)

An average total score of 107.4 out of 133 ($SD=8.6$) was achieved for the PQ. Detailed ratings for each subscale are shown in Table 1. The further calculation revealed a moderate negative correlation between self-reported stress and PQ score. The mean critical gap for the crossing cases also had a moderate negative correlation with the presence, whereas a weak positive correlation was observed for the stepping cases.

4 Discussion

4.1 Effects of the two crossing methods (CMs) on pedestrian behavior

The most compelling finding of the experiment was that the variable of the CM indeed affected participants’ crossing behavior in terms of both critical gap size and subjectively reported stress level. On average, the critical gap was more than double when participants actually crossed the road compared to taking single steps. Also, the subjectively reported stress levels rose moderately in the crossing cases compared to stepping. We suspect that elevated stress levels are an indicator of higher perceived risk, which has been shown to influence participants’ decision-making (Kwon et al. 2022; Papadimitriou et al. 2017).

Fig. 5 Box plots comparing the mean of both dependent variables for “inattentive drivers” in this study and a reference study
Abb. 5 Boxplots zum Vergleich der beiden abhängigen Variablen für unaufmerksame Fahrer mit einer Referenzstudie

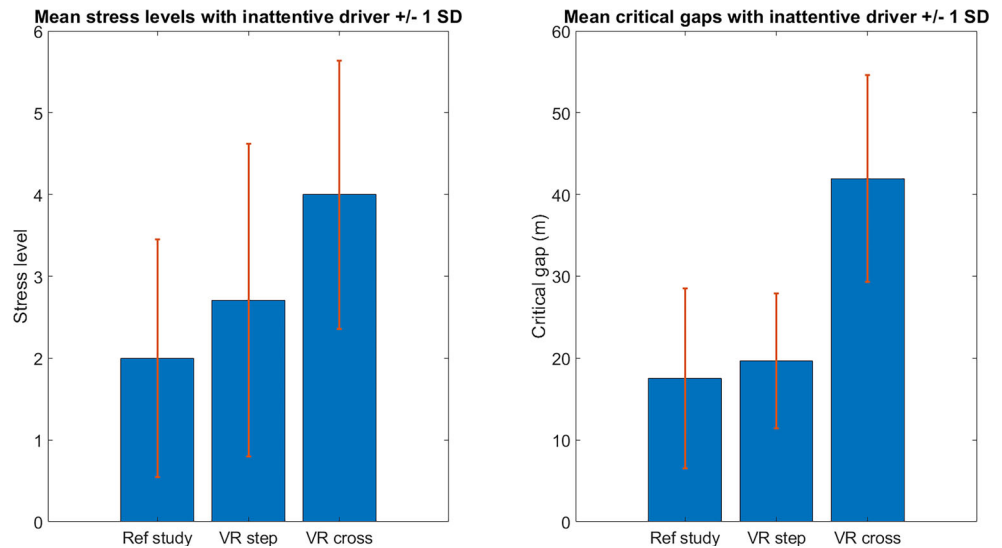


Figure 5 presents a comparison of the results obtained in our study with those reported in the reference study (Rodríguez Palmeiro et al. 2018), considering the most similar experimental conditions. Specifically, we averaged the outcomes obtained with the “inattentive driver” and compared them with the corresponding values from the reference study. We found that the measured critical gap and stress levels during stepping interactions were largely consistent with the real-world data. Nevertheless, we acknowledge that further research is needed to fully validate our findings and explore potential differences between VR-based experiments and real-world field studies.

Most participants reported having made their crossing/stepping decisions based on the speed of the vehicle and the distance from the vehicle, which is in line with the findings of both Rodríguez Palmeiro et al. (2018) and other previous research (Liu and Tung 2014; Sucha et al. 2017). The yielding pattern of the vehicle and the acoustic cues were the next most frequently mentioned factors. Surprisingly, the majority of the participants did not seem to have taken into account eye contact with the driver or his attentional state, which is also consistent with Rodríguez Palmeiro et al. (2018). However, it is important to note that the focus of this work was not to study the effects of explicit eHMI. Thus, no variations in vehicle appearances were included, which limited the selection. Generally, our results broadly support the work of existing field observations, confirming the importance of the vehicle’s implicit eHMI, i.e., its motion patterns, as cues for pedestrians’ decision-making (Dey and Terken 2017; Moore et al. 2019; Rothenbücher et al. 2016). All in all, most of the data recorded by our VR simulator are in line with previous studies, and the experience of moving in front of an operating vehicle could be conveyed without any safety risks.

4.2 Evaluation of the methodology

Our paradigm received an average score of 6.9 for the overall level of realism, which is slightly higher than the score of 6.4 reported by Rodríguez Palmeiro et al. (2018). By contrast, only three participants thought that the VR environment was dissimilar to a real-life situation, which supports the ecological validity of our setup. The combination of visual and acoustic cues was perceived as an advantage by many, while the lack of other pedestrians and drivers and the last-moment constraint for the CMs were criticized. Additionally, the PQ received over 75% of the scores for all subscales (ref. Table 1). We thus conclude that a reasonable level of presence and immersion was achieved by our simulation.

Additionally, neither CM was rated as a very natural interaction, and a moderate correlation ($r=0.4$) between naturalness ratings for the two CMs was observed. Interestingly, most ratings for the naturalness were polarized, and the participants tended to give both high (8–10) and low (0–3) scores. According to the reasons reported by the participants who gave low ratings, defining the last acceptable moment at which to cross or step back was difficult and counterintuitive. This secondary task of determining the right moment may have generated an extra cognitive load for the participants, which has been shown to affect walking behavior both in VR (Kannape et al. 2014) and in the real world (Springer et al. 2006). Although the last-moment constraint made the calculation of the critical gap efficient, alternative methods should be considered for more natural interactions, some of which will be discussed in Sect. 4.4.

However, as the same conditions applied for both CMs, it is reasonable to assume that the last-moment constraint was not responsible for the significant differences in the dependent variables. The wired connection of the HMD may

have also negatively affected the experience and, therefore, the naturalness rating, as it may have distracted the users (Davis et al. 2014; Slater and Steed 2000). Combined with the results from the PQ, we can conclude that although VR is capable of recreating an immersive environment, the specific design choices for the experiment negatively impacted the overall rating.

4.3 Manipulation check

The questions assessing the effectiveness of our manipulation techniques yielded mixed results. Though more than half of the participants realized the pre-recorded nature of the videos showing the driver's face, only three reported being influenced by this realization when crossing the road. This shows that our attempt to recreate a simple yet convincing representation of the driver fell short.

By contrast, only half of the participants noticed the fact that the vehicle was not manually controlled, and none were affected by this realization. Judging from the other data and participants' subjective arguments, we can conclude that the impact of the manipulations was limited. Controlling the vehicle through computer scripts seems to be a valid technique to effectively reduce the complexity of the setup and is recommended for consideration in future studies.

4.4 Limitations and recommendations for future work

Firstly, like the majority of studies conducted in an academic context, we had a limited group of participants from which to sample. Since almost all the participants were students or staff at ETH Zurich, they are not necessarily representative of the larger population. We recommend future studies consider larger and more diverse sample sizes. Recruiting participants from different backgrounds and age groups may also help to improve the generalizability of the results.

Secondly, we limited the number of active traffic users in the VE to focus on studying the effects of the independent variables in isolation, excluding possible interference. In reality, pedestrians will often face complex crossing situations in mixed traffic settings. Additionally, the environmental factors remained the same for all interactions in our study, which could be said to oversimplify real-life situations. Therefore, to obtain a more holistic understanding of pedestrian crossing behavior, further experiments with more complex traffic situations and more variations in the environmental factors should be conducted.

Thirdly, the last-moment constraint enforced for both CMs was shown to have negatively impacted the perceived naturalness of both CMs. Though it indeed simplified the estimation of the critical gap, this came at the cost of re-

duced naturalness. Given this negative impact, we recommend that further studies find other means for the estimation of the critical gap which do not interfere with pedestrians' behavior. While requiring a greater number of measurements, well-established probabilistic (Tian et al. 1999) or deterministic (Ashworth 1970) approaches may yield better user experiences for the participants. Alternatively, one may try to conceal the last-moment constraints through gamification so that it is a more natural part of the interaction.

In this study, stress levels were measured through subjective reports, as we wanted to adopt existing research methods. However, for future studies, objective measures such as heart rate variability, cortisol levels, and skin conductance should be considered to provide a more accurate and reliable assessment of stress. Additionally, it is important to differentiate between physical and mental stress as the two can have different physiological and psychological effects. In this study, we used the term "stress level" as a whole, which may have confounded the results. For example, the act of crossing in virtual reality may be mentally demanding, leading to an increase in perceived stress but not necessarily a corresponding increase in physical stress. Therefore, future research should carefully consider the measurement of stress and differentiate between physical and mental stress to better understand the effects of stress on individuals.

Last but not least, this study is the first attempt to reveal the possible effects of the lack of a physical crossing. Therefore, we selected only one reference method (stepping) and a limited number of dependent variables. Detailed exploration, including more reference methods and more pedestrian behavior data with other criteria should be conducted for a more systematic understanding of this issue.

5 Conclusion

Physically crossing the street and walking in front of oncoming vehicles are rarely permitted for AV-related controlled experiments, and so their effect on pedestrian behavior has remained unexplored. Our study reveals that this lack of a physical crossing can lead to a significantly lower measured critical gap and perceived stress levels in a similar environment in VR. Our other findings demonstrated consistency with prior real-world studies, indicating the capability of VR experiments in substituting risky physical environments and yielding transferable results for real-world problems. These findings provide a strong indication that pedestrian behaviors measured through indirect means may vary according to the method adopted, which should be taken into account for future studies. An arguable weakness of the study was the contrived nature of the last-moment constraints. We thus recommend that future studies

integrate smarter solutions, such as gamification, to make the interaction feel more natural while retaining the ease of critical gap estimation.

Funding Open access funding provided by Swiss Federal Institute of Technology Zurich

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- 255 Pixel Studios (2020) CITY Package. <https://assetstore.unity.com/packages/3d/environments/urban/city-package-107224>. Accessed 25 Sept 2021
- Ashworth R (1970) The analysis and interpretation of gap acceptance data. *Transp Sci* 4(3):270–280. <https://doi.org/10.1287/trsc.4.3.270>
- Bhagavathula R, Williams B, Owens J, Gibbons R (2018) The reality of virtual reality: a comparison of pedestrian behavior in real and virtual environments. *Proc Hum Factors Ergon Soc Annu Meet* 62(1):2056–2060. <https://doi.org/10.1177/1541931218621464>
- Black Starling Productions (2021) Animated trees package. <https://assetstore.unity.com/packages/3d/vegetation/trees/animated-trees-package-169445>. Accessed 9 Oct 2021
- Brewer MA, Fitzpatrick K, Whitacre JA, Lord D (2006) Exploration of pedestrian gap-acceptance behavior at selected locations. *Transp Res Rec* 1982(1):132–140. <https://doi.org/10.1177/0361198106198200117>
- Curran R, Park SY, Domingo L, Garcia-Mancilla J, Santana-Mancilla PC, Gonzalez VM, Ju W (2018) ¡Vamos! Observations of pedestrian interactions with driverless cars in Mexico. *Automotive'18: Proceedings of the 10th Acm International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp 210–220 <https://doi.org/10.1145/3239060.3241680>
- Davis S, Nesbitt K, Nalivaiko E (2014) A systematic review of cybersickness. *Proceedings of the 2014 Conference on Interactive Entertainment*, pp 1–9 <https://doi.org/10.1145/2677758.2677780>
- Deb S, Carruth DW, Sween R, Strawderman L, Garrison TM (2017) Efficacy of virtual reality in pedestrian safety research. *Appl Ergon* 65:449–460. <https://doi.org/10.1016/j.apergo.2017.03.007>
- Deb S, Strawderman LJ, Carruth DW (2018) Investigating pedestrian suggestions for external features on fully autonomous vehicles: a virtual reality experiment. *Transp Res Part F Traffic Psychol Behav* 59:135–149. <https://doi.org/10.1016/j.trf.2018.08.016>
- Dey D, Terken J (2017) Pedestrian interaction with vehicles: roles of explicit and implicit communication. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp 109–113 <https://doi.org/10.1145/3122986.3123009>
- Dey D, Matvienko A, Berger M, Pflöging B, Martens M, Terken J (2021) Communicating the intention of an automated vehicle to pedestrians: The contributions of eHMI and vehicle behavior. *Inf Technol* 63(2):123–141. <https://doi.org/10.1515/itit-2020-0025>
- Edy (2020) Edy's vehicle physics. <https://assetstore.unity.com/packages/tools/physics/edy-s-vehicle-physics-403>. Accessed 11 Nov 2021
- Equilibrium A (2022) Street New York. <https://assetstore.unity.com/packages/3d/environments/urban/street-new-york-183319>. Accessed 5 Oct 2021
- Fagnant DJ, Kockelman K (2015) Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transp Res Part A Policy Pract* 77:167–181. <https://doi.org/10.1016/j.tra.2015.04.003>
- Golding JF (1998) Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain Res Bull* 47(5):507–516. [https://doi.org/10.1016/S0361-9230\(98\)00091-4](https://doi.org/10.1016/S0361-9230(98)00091-4)
- Habibovic A, Andersson J, Nilsson M, Lundgren VM, Nilsson J (2016) Evaluating interactions with non-existing automated vehicles: three Wizard of Oz approaches. 2016 IEEE Intelligent Vehicles Symposium (IV), pp 32–37 <https://doi.org/10.1109/IVS.2016.7535360>
- Highways England (2016) UK road width restrictions FOI. <https://www.gov.uk/government/publications/uk-road-width-restrictions-foi>. Accessed 13 Oct 2022
- Higuera-Trujillo JL, López-Tarruella Maldonado J, Llinares Millán C (2017) Psychological and physiological human responses to simulated and real environments: a comparison between photographs, 360° panoramas, and virtual reality. *Appl Ergon* 65:398–409. <https://doi.org/10.1016/j.apergo.2017.05.006>
- Joisten P, Alexandí E, Drews R, Klassen L, Petersohn P, Pick A, Schwindt S, Abendroth B (2020) Displaying vehicle driving mode—effects on pedestrian behavior and perceived safety. In: Ahram T, Karwowski W, Pickl S, Tairar R (eds) *Human systems engineering and design II*. Springer, Cham, pp 250–256 https://doi.org/10.1007/978-3-030-27928-8_38
- Kannape OA, Barré A, Aminian K, Blanke O (2014) Cognitive loading affects motor awareness and movement kinematics but not locomotor trajectories during goal-directed walking in a virtual reality environment. *Plos One* 9(1):e85560. <https://doi.org/10.1371/journal.pone.0085560>
- Kooijman L, Happee R, de Winter JCF (2019) How do eHMIs affect pedestrians' crossing behavior? A study using a head-mounted display combined with a motion suit. *Information* 10(12):12. <https://doi.org/10.3390/info10120386>
- Kwon J, Kim J, Kim S, Cho G (2022) Pedestrians safety perception and crossing behaviors in narrow urban streets: an experimental study using immersive virtual reality technology. *Accid Anal Prev* 174:106757. <https://doi.org/10.1016/j.aap.2022.106757>
- Liu Y, Tung Y (2014) Risk analysis of pedestrians' road-crossing decisions: effects of age, time gap, time of day, and vehicle speed. *Saf Sci* 63:77–82. <https://doi.org/10.1016/j.ssci.2013.11.002>
- Löcken A, Golling C, Riener A (2019) How should automated vehicles interact with pedestrians?: a comparative analysis of interaction concepts in virtual reality. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp 262–274 <https://doi.org/10.1145/3342197.3344544>
- Mahadevan K, Somanath S, Sharlin E (2018) Communicating awareness and intent in autonomous vehicle-pedestrian interaction. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp 1–12 <https://doi.org/10.1145/3173574.3174003>
- Moore D, Curran R, Strack GE, Sirkin D (2019) The case for implicit external human-machine interfaces for autonomous vehicles. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp 295–307 <https://doi.org/10.1145/3342197.3345320>

- National Research Council (ed) (2010) HCM2010: highway capacity manual. National Academy of Science
- On-Road Automated Driving (ORAD) committee (2016) Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. SAE International. https://doi.org/10.4271/J3016_201609
- Papadimitriou E, Lassarre S, Yannis G (2017) Human factors of pedestrian walking and crossing behaviour. *Transp Res Proc* 25:2002–2015. <https://doi.org/10.1016/j.trpro.2017.05.396>
- Pro 3D Models (2018) Low-poly Civilian vehicle #5. <https://assetstore.unity.com/packages/3d/vehicles/land/low-poly-civilian-vehicle-5-124987>. Accessed 7 Oct 2021
- Rodríguez Palmeiro A, van der Kint S, Vissers L, Farah H, de Winter JCF, Hagenzieker M (2018) Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment. *Transp Res Part F Traffic Psychol Behav* 58:1005–1020. <https://doi.org/10.1016/j.trf.2018.07.020>
- Ropelato S, Menozzi M, Huang M (2022) Hyper-reoriented walking in minimal space. *Virtual Real* 26(3):1009–1017. <https://doi.org/10.1007/s10055-021-00608-0>
- Rothenbücher D, Li J, Sirkin D, Mok B, Ju W (2016) Ghost driver: a field study investigating the interaction between pedestrians and driverless vehicles. 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp 795–802 <https://doi.org/10.1109/ROMAN.2016.7745210>
- Schmidt S, Färber B (2009) Pedestrians at the kerb—Recognising the action intentions of humans. *Transp Res Part F Traffic Psychol Behav* 12(4):300–310. <https://doi.org/10.1016/j.trf.2009.02.003>
- Singh S, Payne SR, Mackrill JB, Jennings PA (2015) Do experiments in the virtual world effectively predict how pedestrians evaluate electric vehicle sounds in the real world? *Transp Res Part F Traffic Psychol Behav* 35:119–131. <https://doi.org/10.1016/j.trf.2015.10.012>
- Slater M, Steed A (2000) A virtual presence counter. *Pres Teleop Virtual Environ* 9(5):413–434. <https://doi.org/10.1162/105474600566925>
- Springer S, Giladi N, Peretz C, Yogeve G, Simon ES, Hausdorff JM (2006) Dual-tasking effects on gait variability: the role of aging, falls, and executive function: Dual-Tasking Effects on Gait Variability. *Mov Disord* 21(7):950–957. <https://doi.org/10.1002/mds.20848>
- Sucha M, Dostal D, Risser R (2017) Pedestrian-driver communication and decision strategies at marked crossings. *Accid Anal Prev* 102:41–50. <https://doi.org/10.1016/j.aap.2017.02.018>
- Tian Z, Vandehey M, Robinson BW, Kittelson W, Kyte M, Troutbeck R, Brilon W, Wu N (1999) Implementing the maximum likelihood methodology to measure a driver's critical gap. *Transp Res Part A Policy Pract* 33(3–4):187–197. [https://doi.org/10.1016/S0965-8564\(98\)00044-5](https://doi.org/10.1016/S0965-8564(98)00044-5)
- U.S. Department of Transportation, Federal Highway Administration (2007) Mitigation strategies for design exceptions. https://safety.fhwa.dot.gov/geometric/pubs/mitigationstrategies/chapter3/3_lanewidth.cfm. Accessed 13 Oct 2022
- Van Brummelen J, O'Brien M, Gruyer D, Najjaran H (2018) Autonomous vehicle perception: the technology of today and tomorrow. *Transp Res Part C Emerg Technol* 89:384–406. <https://doi.org/10.1016/j.trc.2018.02.012>
- Walker F, Dey D, Martens M, Pfleging B, Eggen B, Terken J (2019) Feeling-of-safety slider: measuring pedestrian willingness to cross roads in field interactions with vehicles, pp 1–6 <https://doi.org/10.1145/3290607.3312880>
- Williams B, Narasimham G, Rump B, McNamara TP, Carr TH, Rieser J, Bodenheimer B (2006) Exploring large virtual environments with an HMD on foot. *ACM SIGGRAPH 2006 Research Posters on—SIGGRAPH '06*, p 182 <https://doi.org/10.1145/1179622.1179832>
- Witmer BG, Jerome CJ, Singer MJ (2005) The factor structure of the presence questionnaire. *Pres Teleop Virtual Environ* 14(3):298–312. <https://doi.org/10.1162/105474605323384654>
- Zimmermann R, Wettach R (2017) First step into visceral interaction with autonomous vehicles. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp 58–64 <https://doi.org/10.1145/3122986.3122988>