Interaction Effects of Pedestrian Behavior, Smartphone Distraction and External Communication of Automated Vehicles on Crossing and Gaze Behavior

Mirjam Lanzer
mirjam.lanzer@uni-ulm.de
Human Factors, Ulm University
Ulm, Germany

Ina Koniakowsky
ina.koniakowsky@uni-ulm.de
Human Factors, Ulm University
Ulm, Germany

Mark Colley
mark.colley@uni-ulm.de
Institute of Media Informatics, Ulm University
Ulm, Germany

Martin Baumann
martin.baumann@uni-ulm.de
Human Factors, Ulm University
Ulm, Germany

ABSTRACT

External communication of automated vehicles is proposed to replace driver-pedestrian communication in ambiguous crossing situations. So far, research has focused on simpler scenarios with one attentive pedestrian and one automated vehicle. This virtual reality study (N=115) investigates a more complex scenario with other crossing pedestrians, a distracting task on the smartphone, and external communication by the automated vehicle. Interaction effects were found for crossing duration, gaze behavior, and subjective measures. For attentive pedestrians, the external communication resulted in shorter crossing durations, higher perceived safety, as well as lower perceived criticality, cognitive workload, and effort. These positive effects were not found when pedestrians were distracted. Instead, distracted pedestrians benefited from other crossing pedestrians because they looked less at the stopping vehicle, felt safer, perceived the situation as less critical, and reported lower cognitive workload and effort. Pedestrians initiated crossings earlier with a group or external communication and later with a smartphone.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

automated vehicles, eHMI, virtual reality, pedestrian group, smartphone distraction, unsignalized crossing, eye tracking

ACM Reference Format:
Mirjam Lanzer, Ina Koniakowsky, Mark Colley, and Martin Baumann. 2023. Interaction Effects of Pedestrian Behavior, Smartphone Distraction and External Communication of Automated Vehicles on Crossing and Gaze Behavior. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3544548.3581303

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3581303

1 INTRODUCTION

Automated vehicles (AVs) are expected to change traffic profoundly and bring advantages to various domains such as the environment [86] or safety [32, 44]. In urban areas, up to 73% of analyzed pedestrian crashes would be avoidable with fully AVs [92]. However, AVs also bring new challenges as driver-pedestrian communication is no longer available. To address this challenge, many external Human-Machine Interface (eHMI) concepts were proposed and evaluated for AVs [10, 12, 17, 39]. According to a taxonomy by Dey et al. [23] as well as by Colley and Rukzio [13], they differ, among others, in their communication modality (e.g., lightband or text for visual eHMIs, auditory signals), their placement (e.g., on the vehicle, on the road), their information content (e.g., driving mode, intent) and their degree of scalability (e.g., single or multiple road users).

Previous studies evaluated eHMI concepts almost exclusively in a one pedestrian – one AV scenario [16] where the pedestrian had **no other task** than crossing the street. However, from a large body of observational studies, 38 factors were identified that influence pedestrians' crossing decisions, including characteristics of the pedestrian and the surrounding environment [15, 81]. Observational studies show that pedestrians are rarely alone on the street but are accompanied or surrounded by other pedestrians [61, 73]. Also, pedestrians often handle a secondary task in addition to crossing the street, such as using their smartphone [35, 51, 77, 89, 93]. All these factors that influence pedestrians' crossing behavior do not exist in isolation, but they interact with each other. Understanding how these factors interact with each other is particularly important as traffic is a complex system in which there is practically no situation with only one influencing factor. Hence, the main effects found so far cannot simply be transferred to more complex scenarios, as it seems unlikely that a factor is so strong that it has the same effect in every context. It is, therefore, particularly important to understand under which circumstances factors (e.g., eHMIs) have an effect. While there is an extensive body of literature comparing different eHMI concepts or features (see [5, 23] for an overview), until now, more than one pedestrian [8, 26, 96], pedestrian's attentional state [11] or smartphone distraction [50] have only been evaluated in solitude. In their overview, Rasouli and Tsotsos [81], however, conclude that their interactions have not been sufficiently investigated in research so far.

Studying the interaction of these three factors is especially interesting as all of them either promote or hinder the intake of information that is necessary for the decision to cross the street. The presence of other crossing pedestrians and the eHMIs of AVs can provide cues to the pedestrians to help them correctly assess the criticality of the situation, make a crossing decision and feel safe doing so. Pedestrians' smartphone distraction, on the other hand, interferes with and negatively influences these processes. Thus, it can be expected that these three factors interact with each other. Furthermore, all three factors are not marginal phenomena but are established in research and, in relation to the presence of other pedestrians and smartphone distraction, frequently observed in reality. There are already reliable findings from research on all three variables separately [8, 11, 50]. However, these have never been studied together. There are already voices in the research community calling for these variables to be studied in their interaction [16, 26, 50].

Thus, we conducted a Virtual Reality (VR) study (N = 115) to examine potential interactions between pedestrian factors with the external communication of AVs. Participants crossed the street while three factors were varied: (i) they were distracted by using a smartphone, (ii) another group of pedestrians was also crossing, and (iii) the AVs were communicating their intent via light band eHMIs. Interaction effects were found, especially in relation to smartphone distraction. Being distracted or attentive seems to be a key determinant of pedestrians' crossing and gaze behavior that interacts with the presence of other crossing pedestrians and the external communication of AVs. When pedestrians were attentive, eHMIs led to a decrease in crossing duration, perceived criticality, cognitive workload, and effort and an increase in perceived safety. However, these positive effects were not found when pedestrians were distracted by using a smartphone. The presence of another group of simulated pedestrians crossing made distracted pedestrians feel safer, they perceived the situation as less critical, reduced their cognitive workload and effort. The pedestrians also looked more at the stopping AV when the eHMI was active but only when not distracted by a smartphone and no other pedestrians were there. When other pedestrians were there, they looked less at the stopping AV. In addition, pedestrians initiated their crossing sooner when the eHMIs were active, when other pedestrians were crossing, and later when they were distracted by a smartphone. They looked less at traffic when they were distracted. They also performed worse in the smartphone task when the eHMIs were active.

Contribution statement: This work extends eHMI research by including the pedestrian factors smartphone distraction and other persons' behavior and investigating their interaction with the external communication of AVs. The results of a VR experiment with N = 115 participants showed that the positive effects that eHMIs had on attentive pedestrians disappeared as soon as the pedestrians were distracted by using a smartphone. In contrast, the presence of other crossing pedestrians appeared to benefit distracted pedestrians. This work highlights the need for more complex traffic scenarios as well as the investigation of interaction effects in pedestrian-AV interaction research.

2 RELATED WORK

This work builds on prior work on external communication of AVs, pedestrians distracted by their smartphones, and the presence of other pedestrians. According to Rasouli and Tsotsos [81] and Colley et al. [15], these factors can be categorised as environmental factors (eHMI as a vehicle factor) and pedestrian factors (smartphone distraction as an attentional state factor, presence of other pedestrians as a social factor). Related work on each of these factors will be provided in the following.

2.1 External Communication of Automated Vehicles

External HMIs are designed to help interpret whether a vehicle will yield to a pedestrian [19, 25, 47, 58] when driver-pedestrian communication is no longer present in AVs. While some researchers argue that implicit communication, such as the vehicle's movement patterns or engine sound, is sufficient [27, 65, 72], the majority of studies seem to find positive effects of eHMIs. When AVs are equipped with eHMIs, pedestrians feel safer to cross [19, 31, 64], trust the AV more [53, 70], have a reduced cognitive load [8], and initiate their crossing sooner [31, 47, 58]. de Winter and Dodou [20] provide an overview of the arguments for and against the necessity of eHMIs.

Most of these studies have been done in a simplified scenario with one attentive pedestrian and one AV. As this is a logical first step when investigating a new technology, the focus of current research in the eHMI domain seems to shift towards topics that take pedestrian factors into account. This includes studies about different cultural backgrounds [53, 60], age groups [21, 45], and including pedestrians with impairments [4, 14, 39].

Understanding pedestrians' attentional state is crucial for pedestrian safety in automated traffic, as pedestrians' head orientation and gaze behavior are indicators that AVs take into account when predicting pedestrians' intention to cross the road [57]. Distracted pedestrians in particular show a less pronounced gaze and head orientation pattern towards traffic than attentive pedestrians (see subsubsection 2.2.1) which must nevertheless be correctly interpreted by AVs in order to avoid accidents.

Colley et al. [11] investigated whether pedestrians who were distracted by a cognitively demanding task on a billboard in front of them could benefit from visualizing the direction of oncoming traffic on parked vehicles or the pavement but did not find significant interaction effects. Holländer et al. [50] found that guidance given on the smartphone screen helped distracted pedestrians make successful crossing decisions in non-automated traffic and reduced their cognitive workload. To the authors' knowledge, smartphone distraction and external communication of AVs have not yet been studied together.

Addressing the issue of scalability of eHMIs, AVs should provide clear and unambiguous communication, such as displaying the yielding intention to avoid miscommunication when multiple pedestrians are present [26, 96]. Colley et al. [8] investigated the effect of eHMIs while a group of pedestrians performed a critical crossing in front of a non-yielding AV. They found that most participants did not directly follow the group but still crossed earlier than when no group was present. In a video study [53], pedestrians stated that they felt safer in automated traffic when waiting at the curb with a group of simulated pedestrians, but no effects were found for the crossing willingness. The authors suggest that this could be due to the imitation effect not occurring, as the group of simulated pedestrians were only waiting but not crossing, which is also supported by Colley et al. [9]. In both studies, either no interaction effects were calculated or detected.

2.2 Pedestrian Factors

2.2.1 Smartphone Distraction. For a safe crossing, it is crucial that pedestrians are attentive before and while crossing the street [82]. Yet, observational studies conducted in multiple countries and across different traffic environments report numbers between 20-30% of pedestrians using their smartphones while crossing [35, 51, 77, 89, 93]. For pedestrians, visual, cognitive and auditory distraction pose a safety risk [43]. Visual distraction can be caused by any task that involves perceptual processes that limit a pedestrian's visual abilities [2]. Looking at a smartphone can constitute a visual distraction as pedestrians may not see approaching vehicles or relevant traffic signs. Cognitive distraction refers to any task that requires cognitive processing and results in thinking about something unrelated to the crossing task [63]. According to Strayer et al. [87], cognitive distraction occurs when the cognitive workload exceeds a certain threshold. This can happen if a competing task (e.g. texting on a smartphone) is performed at the same time as crossing the road, as both require cognitive resources (e.g. estimating the speed of oncoming vehicles). A systematic review across 14 studies found that among all smartphone-related activities, texting, a visually and cognitively distracting task [43], has the most detrimental effects on pedestrians' behavior [85]. When being distracted by a smartphone, pedestrians feel less safe [98] and report a higher perceived workload and an impaired situational awareness [67]. Being distracted while walking is a predictor of unsafe crossing, meaning pedestrians are less likely to follow a straight path and cross more slowly [43, 66], which increases their risk for collisions [74, 82]. They also take longer to initiate a crossing when there is a safe gap between vehicles [6, 85]. In addition, pedestrians distracted by a smartphone look less at oncoming traffic [43, 85, 93], reduce their scanning frequency, and fixate less and shorter on traffic-related areas [52].

2.2.2 Effects of Other Pedestrians' Behavior. Observational studies showed that 50-70% [34, 73] and, in the city center, up to 88% [61] of pedestrians cross in groups, with the most commonly observed group size being two to three people [61]. Even if they are travelling alone, they are usually surrounded by other pedestrians. Pedestrians report feeling safer when crossing in a bigger group [98]. This also influences the pedestrians' crossing behavior and the drivers' behavior. At unsignalized crossings, drivers are likelier to yield to pedestrian groups of three or more people [88]. On the other hand, pedestrians are more likely to cross [33, 80] and to cross faster [8] if other pedestrians have already started to cross, compared to when no other pedestrians are present. Also, individual pedestrians are less careful when crossing a road if several people cross at the same time [41] and less likely to check for oncoming traffic [61, 76]. Instead, pedestrians are more likely to look towards pedestrians when crossing in the presence of others [28]. Following the behavior of other pedestrians may be an indicator of compensatory behavior [85], meaning that relying on groups' social information when to cross the street spares one's own resources [40].

3 PEDESTRIAN STUDY

To investigate interaction effects between factors influencing pedestrians' crossing and gaze behavior, we used a 2 x 2 x 2 within-subjects design. The independent variable on the vehicle side was whether the *eHMI*s of the AVs were activated. The independent variables on the pedestrian side were whether there was another *group* of pedestrians crossing and whether the pedestrian was distracted by using a *smartphone*. As we used a full factorial design, each participant experienced eight conditions. The order of the conditions was randomized using a Latin square. Hypotheses for the main effects could be derived from the literature. However, the main focus of this study was to examine potential interaction effects. As this is, to the authors' knowledge, the first paper to examine the interaction of these variables, any potential interaction effects were exploratory in nature. The hypotheses and research questions of this study are shown in Table 1.

Table 1: Hypotheses (H) and Research Questions (RQ) of the study

	Dependent variable	Independent variable	Hypothesis
H1a	Crossing initiation time	eHMI	When the AV communicates via an *eHMI* compared to when not, pedestrians **initiate** their **crossing sooner**.
H1b		Group of pedestrians	When a *group of pedestrians* is crossing in front compared to when not, pedestrians **initiate** their **crossing sooner**.
H1c		Smartphone distraction	When pedestrians are *distracted by a smartphone* compared to when not, they **initiate** their **crossing later**.
H2a	Crossing duration	Group of pedestrians	When a *group of pedestrians* is crossing in front compared to when not, pedestrians have a **shorter crossing duration**.
H2b		Smartphone distraction	When pedestrians are *distracted by a smartphone* compared to when not, they have a **longer crossing duration**.
H3a	Gaze behavior towards traffic	eHMI	When the AV communicates via an *eHMI* compared to when not, pedestrians **look less** at **traffic-related areas**.
H3b		Group of pedestrians	When a *group of pedestrians* is crossing in front compared to when not, pedestrians **look less** at **traffic-related areas**.
H3c		Smartphone distraction	When pedestrians are *distracted by a smartphone* compared to when not, they **look less** at **traffic-related areas**.
H4a	Perceived safety	eHMI	When the AV communicates via an *eHMI* compared to when not, pedestrians **feel safer**.
H4b		Group of pedestrians	When a *group of pedestrians* is crossing in front compared to when not, pedestrians **feel safer**.
H4c		Smartphone distraction	When pedestrians are *distracted by a smartphone* compared to when not, they **feel safer**.
H5a	Perceived cognitive workload	eHMI	When the AV communicates via an *eHMI* compared to when not, pedestrians have a **lower cognitive workload**.
H5b		Smartphone distraction	When pedestrians are *distracted by a smartphone* compared to when not, they have a **higher cognitive workload**.
RQ	What interactions with each other do the variables *eHMI* of an AV, presence of a *pedestrian group* and *smartphone distraction* have on crossing initiation time, crossing duration, gaze behavior, secondary performance in the smartphone task, perceived safety, perceived criticality, and perceived workload?		

3.1 Sample

The final sample consisted of 115 people (78 female, 36 male, 1 non-binary). For details on excluded participants (n = 9), see subsubsection 3.5.1. Participants' age ranged from 19 to 61 years with an average age of M = 24.85 years (SD = 7.40 years). The sample consisted of 83% students and 17% employees. Most participants stated to walk daily (67%) or close to daily (on 5-6 days per week; 21%), while the remaining 12% walked on 2-4 days a week. On average, participants walked 31-45 minutes daily, covering a distance of 2-3 km per day. Three participants had been involved in an accident as a pedestrian in the past five years, with minor (n = 2) or no (n = 1) injuries. Most participants (89%) reported being right-handed, while 11% reported being left-handed.

Participants were recruited through flyers, social media, and various mailing lists. They were requested to be at least 18 years old and fluent in German. As the study was conducted in VR, people with motion sickness or members of a risk group (e.g., pregnant women or people with epilepsy) were not allowed to participate in the experiment. All participants had normal or corrected to normal vision, and people with glasses were asked to wear contact lenses during the study. The experiment was approved by the ethics committee of Ulm University.

3.2 Study Setup

3.2.1 Crossing Scenario. The crossing scenario was adapted from Colley et al. [8] and implemented in VR using Unity (version 2020.03. 1f10). A Vive Pro Eye VR setup was used with three base stations covering an area of 6 x 4 m^2 (see Figure 1).

Figure 1: Laboratory setting and VR setup. A participant walking with the Vive Pro Eye VR headset on and a controller in her hand (left). Experimenter perspective and experimenter computer with Unity (right).

When participants entered the simulation, they were standing in a park close to the sidewalk of a two-lane street in an urban area (see Figure 2). They were instructed to cross the street to reach their destination on the opposite side (highlighted in green). Due to space and tracking constraints, the same gain factor of 1.7 that Colley et al. [8] used in their study was added in the straightforward and sideways (not height) axis. This means that when participants cover 1 m in the lab, they cover 1.7 m in the VR scenario. Vehicles were approaching from both directions at a speed of approximately 30 km/h. During the first 28 s, traffic was dense and gaps between vehicles were approximately 3s long. They were considered critical crossing opportunities as they are too short for a pedestrian to safely cross the street. Therefore, it was expected that participants would not cross the street during the first 28 s. After 28 s, a vehicle coming from the left decelerated and stopped in front of the participant (stopping AV). The next vehicle from the right arrived after an additional 14 s, creating an adequate gap in which participants could safely cross the street. The traffic was identical, especially regarding vehicle sequence and gaps, in each scenario to ensure comparability. All virtual vehicles were compact cars, identical in appearance, drove autonomously, and were introduced as driverless cabs.

3.2.2 External HMI. External communication was operationalized by a light strip, displaying the vehicle's intent (driving vs. stopping; see Figure 3). This was adapted from Colley et al. [8] who followed the approach by Faas et al. [31] and Dey et al. [24]. This kind of eHMI was chosen as it is technically feasible (compared to, for example, projections or windscreen displays) and does not require language skills (compared to text) [8]. An intention-based light band eHMI is also visible to multiple pedestrians and has already been studied regarding scalability with positive results in terms of crossing willingness and decision certainty [96]. The eHMI was attached to the lower front of the vehicle: when the vehicle was driving, two outward-moving yellow dots were visible on the LED band. When it slowed down and stopped, the LED strip flashed turquoise.

3.2.3 Other Pedestrians. The influence of other crossing pedestrians was examined by using a group of three virtual pedestrians, consisting of two men and one woman (see Figure 4) The characteristics of the simulated pedestrian group (age, gender, group size) was adapted from Colley et al. [8] and corresponds to typically observed group constellations [61]. The pedestrian group was located at the same side of the street, to the right of the participants. The group was placed slightly set back so that the visibility of traffic was not reduced, but the group was still clearly visible when looking to the right. The simulated pedestrians initiated their crossing 1 s before the vehicle from the left (stopping AV) came to a standstill. This ensured that they crossed the street before the participant.

3.2.4 Smartphone Distraction. The distraction task was implemented by a virtual smartphone on a handheld controller (see Figure 5). On this virtual smartphone, pedestrians performed the n-back task, a cognitive performance task that has been used in several pedestrian studies [11, 90, 95]. As texting on a smartphone, a visually and cognitively demanding task [56], showed the most detrimental effects on pedestrian behavior [85], a visual version of the n-back task [94] seems to be an appropriate way to test smartphone distraction while measuring performance on this secondary task in a standardized way. Thus, a simple visual version (1-back task) was used to prevent cognitive resources from increasing to the point where pedestrian safety was compromised. Participants had to indicate whether the letter displayed matched the previous letter or not via a key press on the controller that they held in their dominant

Figure 2: Bird's eye view of the scenario. (Top) The vehicle coming from the left has already stopped. The next vehicle from the right came at a distance of 14 s. This created a gap in which the participant (red cross) could cross. (Bottom) Four areas (park, sidewalk, street, destination) were used for the logging. The participants started in the park and waited on the sidewalk (red cross) to cross the street. The pedestrian group was waiting at the same height, to the right of the participant on the sidewalk (circled in blue). The sidewalk was 1.8 m, and the street was 6.5 m wide in VR. The destination area was on the sidewalk, on the other side of the street, and was highlighted in green.

Figure 3: The three modes of eHMI that were operationalized by an LED band on the lower front of the vehicle: animated yellow dots represented driving (left), flashing turquoise band represented stopping (center), and no lights represented inactive eHMI (right).

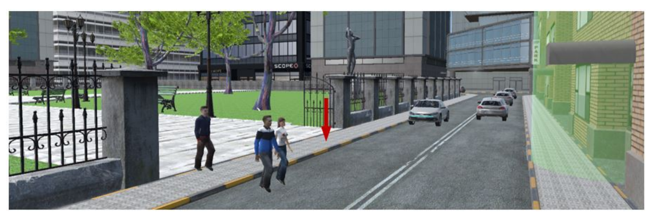

Figure 4: The simulated group of pedestrians started crossing the street shortly before the vehicle came to a halt. The red arrow indicates the location of the participant.

hand. A match was handled correctly by pressing the button on the front with the thumb, and a mismatch was handled correctly by pressing the button on the back of the controller with the index finger. Participants were instructed that correctness and not speed was the goal of this task. The letters were displayed for 2 s each, and participants had the chance to respond within this time. They received immediate feedback on whether the trial had been handled correctly or incorrectly by displaying a green circle or a red cross at the top of the smartphone screen. A blank screen followed for 1 s, resulting in a new letter appearance every 3 s. A different sequence of letters was used in each trial, but the ratio of matches to mismatches was kept the same. The participants started the 1-back task by themselves as soon as they entered the scenario. The task, which was performed continuously throughout the entire duration of the scenario, could not be paused or interrupted by the participants and ended when the experimenter stopped the simulation, i.e., as soon as the participants reached the other side of the street. The length of the task was chosen in such a way that it was not possible to end the task before finishing the crossing process. The participants were instructed to engage with the task while safely crossing the street. On average, they engaged with the smartphone 9 times per condition.

Figure 5: Virtual smartphone with the 1-back task from the pedestrian perspective. A letter is displayed, and the red cross above the letter indicates that the participant had handled the trial incorrectly.

3.3 Measurements

3.3.1 Objective Measurements. The VR scenario was divided into four areas, i.e., park, sidewalk, street, and destination (see Figure 2), and the participants' position was logged within these areas with 50 Hz. The crossing initiation time was determined by the time that it took for the participants to step onto the street after the car coming from the left (stopping AV) had stopped. The crossing duration time was determined by the time that the participants were on the street.

For gaze behavior, five Areas of Interest (AoIs) were defined in advance (traffic, the stopping AV from the left, smartphone, other pedestrians, and destination area). The participants' gaze was automatically detected in these predefined AoIs. Dwell time was calculated, i.e., the time a participant spent looking at an AoI, including all fixations and saccades in an AoI. No gazes shorter than 100 ms were included in the dwell time as the information that can be processed in 100 ms is limited [91]. As blinks were not automatically detected, gaze data was interpolated when a blink occurred while participants looked at an AoI. A blink was defined when participants' gaze was not detected for less than 150 ms, as blink duration is around 100-150 ms on average [7]. The frequency of gazes in an AoI as well as the time participants looked at an AoI (e.g., smartphone) in relation to total time, were calculated as dependent variables [68]. Only gazes, while the participants were standing on the sidewalk prior to crossing, were included in the analyses. This ensured that all participants had a similar field of view and that the gaze reflected the decision-making process when crossing as gazes when walking to the curb are primarily directed at the crossing infrastructure, as opposed to gazes at traffic or other AoIs [36].

For assessing the secondary task performance, the error rate in the 1-back task on the virtual smartphone was computed. Misses and false alarms were treated as errors, and hits and correct rejections were treated as correct trials. The error rate was calculated from all letters processed until the crossing was initiated.

3.3.2 Subjective Measurements. As for subjective measures, the raw NASA TLX scales [42] were used to assess mental demand ("How mentally demanding was the task?"), physical demand ("How physically demanding was the task?"), temporal demand ("How hurried or rushed was the pace of the task?"), performance ("How successful were you in accomplishing what you were asked to do?"), effort ("How hard did you have to work to accomplish your level of performance?"), and frustration ("How insecure, discouraged, irritated, stressed, and annoyed were you?"). Perceived safety ("How safe did you feel in the situation?"; adapted from [46]) and perceived criticality ("How critical did you find the situation you just experienced?"; [62]) were assessed as single items on a 5-point and a 7-point Likert scale respectively (1=not at all, 5/7=completely).

3.3.3 Interview. In the interview, participants answered seven questions about their experience during the experiment. They were asked first what influenced their crossing decision. Afterwards, each of the three factors (*eHMI, smartphone distraction, pedestrian group*) was addressed explicitly and participants had the chance to elaborate in which way this factor influenced their crossing decision. Participants were also asked whether they were particularly tired or unconcentrated and whether they were able to show the crossing behavior that they would normally show in real traffic. Finally, they had the opportunity to make general comments about the study.

3.4 Procedure

An overview of the procedure can be seen in Figure 6. The experiment was conducted in a laboratory setting in compliance with Covid-19 regulations. After being informed about the study subject, participants gave informed written consent. They were told that the study was about pedestrians' crossing behavior but were not given any further details. Participants were instructed to cross the street to reach their destination on the opposite side of the street. They were told to behave as they would in real traffic and that there was no time pressure to cross the street particularly fast but rather to cross safely. An introduction to the VR setup followed, along with practice trials. In a short familiarization phase, participants got used to the environment without any traffic and finished their task of crossing the street by walking through the lab. As soon as they reached the destination area on the other side, the VR simulation was stopped. Then, the traffic and the group of crossing pedestrians were introduced. It was explained that all vehicles were driving fully automated without a driver being present. No details were given about the participant's relation to the other group of pedestrians (e.g., friends, strangers). The communication via the eHMI was explained, shown in a video, and also experienced in a test trial in VR. Lastly, the smartphone task was instructed, and participants practiced it in VR. Participants were able to ask questions at any given time. When participants felt comfortable with the task, the experimental trials started. Participants were instructed that there would always be traffic in these trials and that they were informed beforehand whether they would have to do the smartphone task. Then, the eight experimental trials followed. After each trial, participants answered questions about the experienced situation on a tablet (interim questionnaire with the subjective measurements). After completing all trials, they answered questions in a short 10-minute interview and filled in a questionnaire about demographics,

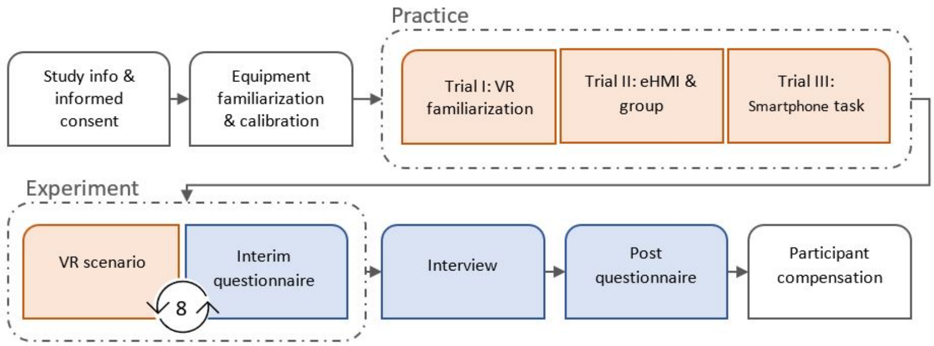

Figure 6: Study Procedure. Orange boxes indicate parts in VR and blue boxes indicate self-report parts. After informed consent was given, participants completed three practice trials. Then eight experimental trials followed, comprising the crossing in the VR scenario and an interim questionnaire. After completing all trials, participants were interviewed and filled in a questionnaire.

immersion [59] and presence [84] in VR, and their pedestrian behavior [38] and walking habits. At the end, participants were debriefed and compensated. Overall, the study took approximately 70 minutes, and participants were compensated with 15€ or course credit. The data collection took place in two periods (May and October 2022).

3.5 Data Preparation and Analysis Procedure

3.5.1 Data Preparation. Two participants were excluded because of technical problems with the VR or the questionnaire and one because of fatigue. Furthermore, 15 participants crossed the street in an earlier gap than the one created by the stopping vehicle from the left. For nine participants, this happened only once (n = 8) or twice (n = 1), so these trials were repeated, and the repetitions were used in the analysis. For six participants, this happened in three or more trials, although they were made aware of this. As these crossings were considered to be critical, these participants were excluded from the data analysis. In n = 13 trials, initiation times were negative because participants mistakenly stood on the street while waiting to cross. This happened unsystematically for different conditions. These trials were listwise deleted. Another n = 4 trials were excluded due to an experimenter's mistake during data collection. Even though eye tracking was used during the entire study, only the smartphone could be reliably detected during the first data collection period due to technical issues. Thus, only the gaze behavior of participants in the second data collection period (n = 59) was analyzed.

Outliers were defined on the trial level for each condition and dependent variable. Data points that were more than 1.5 times the interquartile range were considered to be slight outliers and were not removed as they can be attributed to the manipulation. There were extreme outliers for the crossing initiation time (n = 10 trials), for the crossing duration (n = 3 trials), the secondary task performance (n = 1 trial), and the gaze behavior (n = 9 trials) of more than three times the interquartile range, which were excluded from the analysis since it cannot be ensured that these values were attributed to the manipulation. These trials were pairwise deleted. For the subjective dependent variables, outliers were not influential and thus kept in the analyses.

3.5.2 Analysis Procedure. For the analyses, R (version 4.2.1) and RStudio (version 2022.07.01) were used. Overall, 913 trials were analyzed. As this was a within-subjects design, the eight trials (Level 1) were nested within participants (Level 2). The interdependencies between participant observations were calculated using intraclass correlation (ICC). For all dependent variables, the ICC was > 0.05. Thus, hierarchical linear regressions with a random intercept and a fixed slope were calculated for each dependent variable. All independent variables, i.e., *smartphone distraction* (yes vs. no), *external communication* (yes vs. no), and *crossing group* (yes vs. no), as well as their interactions, were included. The predictors were all effect coded, thus allowing an interpretation of main and interaction effects that is similar to the more commonly used repeated-measures ANOVA. In case of significant interactions, regression models were calculated for the relevant subsets of the data (e.g., with and without the *smartphone distraction*), analogous to the procedure for calculating simple main effects in the ANOVA procedure.

3.5.3 Effects of Time and Data Collection Period. As data collection took place during two separate time periods, this was included as an effect-coded predictor in the models. However, there were no significant effects, i.e., differences, for the two different data collection periods except for initiation time, where participants in the second data collection were generally slightly slower (0.14s on average) to initiate a crossing.

Even though the sequence of conditions was balanced using a Latin square to reduce order effects, time was included in the analyses as a random slope predictor to account for potential individual learning effects. Significant effects were found for the subjective data and the secondary task performance. For perceived safety, perceived criticality, perceived mental and temporal demand, perceived effort, and frustration, the first trial was significantly different from the remaining seven trials (lower for perceived safety and higher for all other variables). Thus, the first trials for these variables were excluded from the analyses as they did not depict the effects of the manipulated variables but were rather an effect of the first contact with the situation. In consequence, excluding them eliminated the effect of time in the final regression models. As the sequence of conditions was balanced, the exclusion did not affect the overall frequency of conditions in the final model. For perceived performance and secondary task performance, there was a significant time effect over all trials. Over time, the participants improved at the smartphone task (fewer errors), which was also reflected in the perceived performance measure. However, as the order of conditions was balanced, this did not systematically influence the effects of the other manipulated variables. No time effects were found for the remaining objective measurements, initiation time, crossing duration, all gaze parameters as well as perceived physical demand. Even though the participants experienced the traffic scenario eight times, this did not alter their crossing and gaze behavior over time. The plots regarding time effects and the complete regression models can be found in the supplementary materials.

4 RESULTS

In our sample, immersion in VR can be considered as high (M = 21.33, SD = 4.35) since it was one standard deviation higher than the reference values provided for young adults (19 - 32 years) (M = 15.87, SD = 5.93; [59]).

4.1 Crossing Initiation Time

The crossing initiation time was defined as the time it took participants to step onto the street after the oncoming vehicle from the left had stopped. Smaller values indicate that participants initiated their crossing sooner and larger values that they initiated the crossing later. Depending on the participant and the condition, crossing initiation time ranged from Min = 0.38 s to Max = 5.36 s. On average, participants were fastest in the condition with a *crossing group* and an *eHMI* (M = 1.74 s, SD = 0.49 s) and slowest in the *smartphone*-only condition (M = 2.31 s, SD = 0.83 s) to initiate a crossing (see Figure 7 left).

As for inferential analyses, neither the three-way nor any two-way interactions were significant. However, a significant main effect for the *crossing group* (β = -0.09, t = -4.96, p < .001), the *smartphone distraction* (β = 0.18, t = 9.90, p < .001) and the *eHMI* (β = -0.04, t = -2.47, p = .014) was found. Participants initiated their crossing earlier when a group was crossing in front and when the eHMI was active. When pedestrians were distracted by a smartphone, they initiated their crossing later.

4.2 Crossing Duration

The crossing duration was defined as the time participants spent on the street. Smaller values indicate that participants crossed faster and larger values that they crossed slower. Depending on the participant and the condition, crossing duration ranged from Min = 1.92 s to Max = 7.58 s. On average, the participants crossed fastest in the *eHMI*-only condition (M = 2.83 s, SD = 0.54 s) and slowest in the condition with *smartphone distraction* and *crossing group* (M = 3.29 s, SD = 0.77 s) (see Figure 7 right).

Inferential analyses revealed a significant three-way interaction (β = -0.03, t = -2.07, p = .039). To examine the interaction, the data was first split by *group*. When there was **no group crossing**, a significant interaction effect of *smartphone distraction* and *eHMI* was found (β = 0.07, t = 2.44, p = .015). When there was **no crossing group** and participants were **not distracted by a smartphone**, the crossing duration was shorter when the *eHMI* was active (β = -0.05, t = -2.68, p = .009). When there was **no crossing group** and participants were **distracted by a smartphone**, the *eHMI* had no effect. When there **was a group crossing**, a main effect of *smartphone distraction* was found (β = 0.20, t = 12.39, p < .001), the *eHMI* had no influence anymore. The participants crossed slower when there was a group and they were distracted by a smartphone.

Second, the data was split by *smartphone distraction*. When participants were **not distracted a smartphone**, a significant main effect of *eHMI* was found (β = -0.03, t = -2.66, p = .008). Participants crossed faster when the eHMI was active compared to when not. When the participants were **distracted by a smartphone**, there

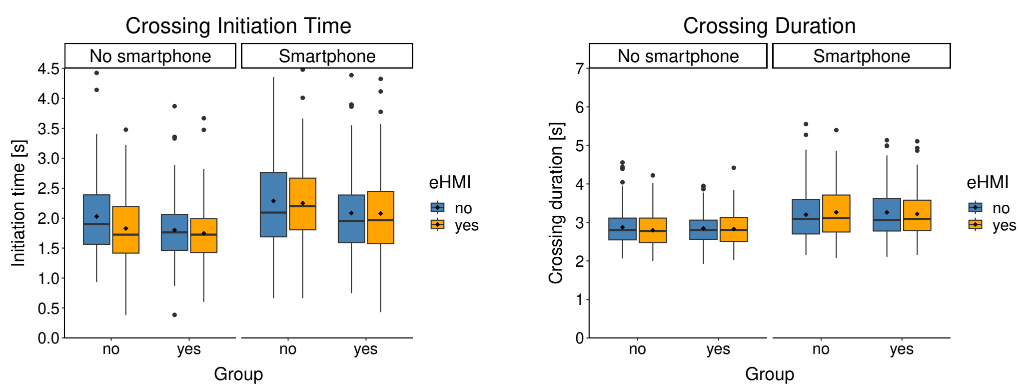

Figure 7: Crossing initiation time (left) and crossing duration (right) by *eHMI* communication, *crossing group*, and *smartphone distraction*. Smaller values represent shorter initiation times/crossing duration. Shorter initiation times with *group* and *eHMI* and longer initiation times with *smartphone distraction*. Three-way interaction for crossing duration. For the details of significance tests, see the text.

was a significant two-way interaction of *eHMI* and *group* (β = -0.05, t = -2.01, p = .045). However, none of the pairwise comparisons between conditions were significant.

Third, the data was split by *eHMI*. In both cases, with and without eHMI, *smartphone distraction* had a significant main effect (**with eHMI**: β = 0.24, t = 8.77, p < .001; **without eHMI**: β = 0.18, t = 10.26, p < .001). The participants crossed slower when they were distracted by a smartphone.

4.3 Gaze Behavior

Gaze behavior was operationalized via AoIs and calculated by the frequency of gazes at an AoI and the time that participants spent looking at the AoI relative to the rest. Smaller values indicate that participants looked less often and shorter at the AoI.

4.3.1 Gazes at Traffic. Depending on the participant and the condition, the frequency of gazes at traffic ranged from *Min* = 0 to *Max* = 34 and the proportional dwell time at traffic ranged from from *Min* = 0.00 to *Max* = 0.78. On average, participants looked at traffic least often in the *smartphone*-only (*M* = 10.28, *SD* = 4.91) and most often in the *eHMI*-only condition (*M* = 19.38, *SD* = 5.09). On average, participants looked proportionally the least long at traffic in the condition with a *smartphone distraction* and a *crossing group* (*M* = 0.18, *SD* = 0.09) and the longest in the condition without *smartphone distraction*, *eHMI* and *group* (*M* = 0.54, *SD* = 0.11) (see Figure 8.1 and 2).

There was a significant effect of *smartphone distraction* on the frequency (β = -4.13, t = -20.56, p < .001) of gazes at traffic. When participants were distracted by a smartphone, they looked less often at traffic compared to when they were not distracted by a smartphone. A significant two-way interaction between *smartphone distraction* and *eHMI* was found for the proportion of gazes on overall traffic (β = 0.01, t = 2.07, p = .040). When participants were **distracted by a smartphone**, the proportion of gazes at traffic was higher when the *eHMI* was active compared to when not (β = 0.01, t = 2.10, p = .040). When participants were **attentive**, no effect of the *eHMI* was found.

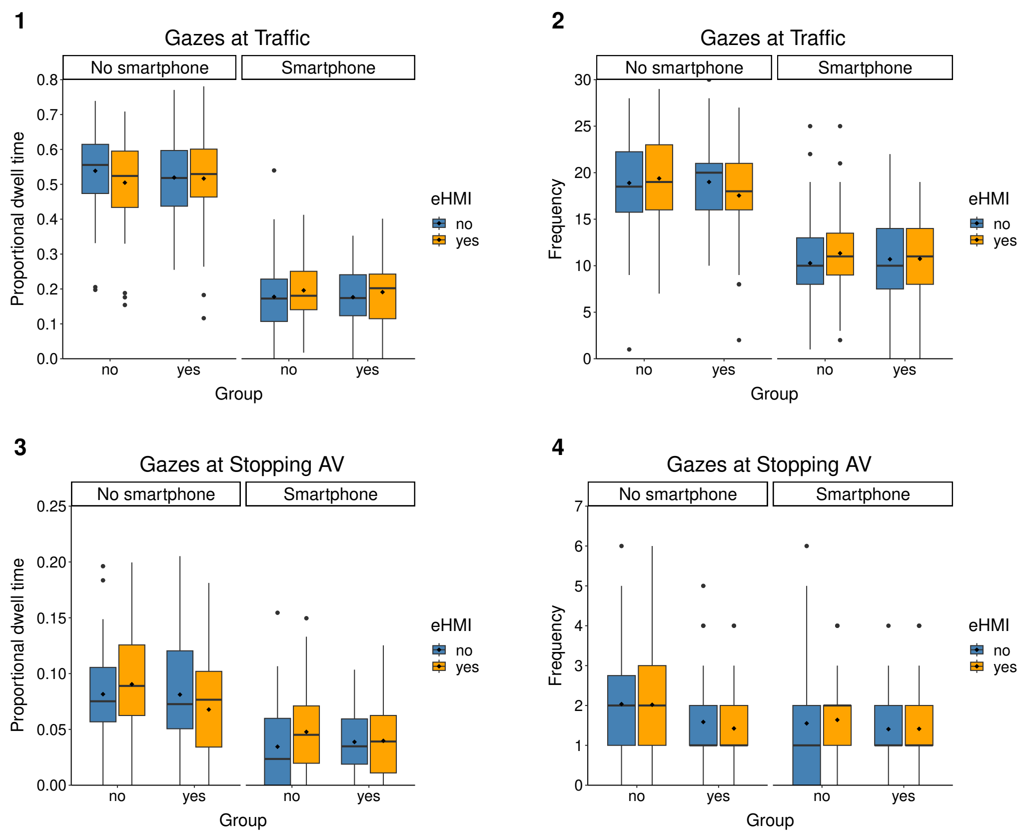

Figure 8: Gazes at traffic (1: proportional dwell time, 2: gaze frequency) and gazes at stopping AV (3: proportional dwell time, 4: gaze frequency) by *eHMI* communication, *crossing group*, and *smartphone distraction*. Smaller values represent shorter proportional dwell time and less gazes at traffic/stopping AV. Two-way interaction of *smartphone distraction* and *eHMI* for proportional dwell time at traffic. Fewer gazes at traffic with *smartphone distraction*. Two-way interaction of *eHMI* and *crossing group* for proportional dwell time at stopping AV. Fewer gazes at stopping AV with *smartphone distraction* and *crossing group*. For the details of the significance tests, see the text.

4.3.2 Gazes at Stopping AV. The frequency of gazes at the stopping AV ranged from *Min* = 0 to *Max* = 6 and the proportional dwell time at the stopping AV ranged from *Min* = 0.00 to *Max* = 0.21, depending on the participant and condition. On average, participants looked at the stopping AV least often in the condition with *smartphone distraction* and *crossing group* (M = 1.41, SD = 0.95) and most often in the condition without *smartphone distraction*, *eHMI* and *crossing group* (M = 2.03, SD = 1.20). On average, participants looked proportionally the least long at the stopping AV in the *smartphone*-only condition (M = 0.03, SD = 0.04) and the longest in the *eHMI*-only condition (M = 0.09, SD = 0.05) (see Figure 8.3 and 4).

A significant main effect of *smartphone distraction* (β = -0.13, t = -2.91, p = .004) and *group* (β = -0.17, t = -3.99, p < .001) was found for the frequency of gazes at the stopping AV from the left. The participants looked less often at the stopping AV when they were distracted by a smartphone compared to when not and when there was a crossing group compared to when not.

For the proportional dwell time of gazes at the stopping AV, there was a significant two-way interaction between *eHMI* and *group* (β = -0.004, t = -2.51, p = .013). When there was **no crossing group**, a higher proportion of gazes towards the stopping AV was observed when the *eHMI* was active compared to not (β = 0.005, t = 2.31, p = .022). When there **was a crossing group**, no influence of the *eHMI* was found. In addition, when the **eHMI was active**, a lower proportion of gazes at the stopping AV was observed when the *group* was crossing (β = -0.008, t = 3.20, p = .002). When the **eHMI was not active**, no influence of the crossing *group* was found. There was also a main effect of *smartphone distraction* (β = -0.02, t = -11.80, p < .001). The participants looked proportionally less long at the stopping AV when distracted by a smartphone.

4.3.3 Gazes at Other Pedestrians and Smartphone. On average, the participants looked 13.96 times at the smartphone (ranging from 1 to 29) and 2.48 (ranging from 0 to 10) times at the other pedestrians. Participants spent an average of 39.58% (ranging from 1% to 79%) of their time looking at the smartphone and 3.27% (ranging from 0% to 25%) of their time looking at other pedestrians, across all conditions where the *smartphone distraction* or *group* was present. Additional information on gazes at other pedestrians and the smartphone can be found in the supplementary materials.

4.4 Secondary Task Performance

The secondary task performance was described as the proportional error rate, i.e., the ratio between the errors in the n-back task and the total number of items completed by participants on the virtual smartphone. Higher values indicate that participants performed worse in the secondary task and made more errors. On average, the participants completed eight items in each condition (ranging from *Min* = 7.78 items to *Max* = 7.94 items). The proportional error rate ranged from *Min* = 0 (no errors) to *Max* = 0.75 (75% errors) depending on the participant and the condition. On average, participants made the most errors in the condition with the *eHMI* (M = 0.18, SD = 0.17) and the fewest errors in the condition with the *crossing group* (M = 0.11, SD = 0.13; see Figure 9).

Inferential analyses found no significant interaction effect but a significant main effect for *eHMI* (β = 0.01, t = 2.44, p = .015). When

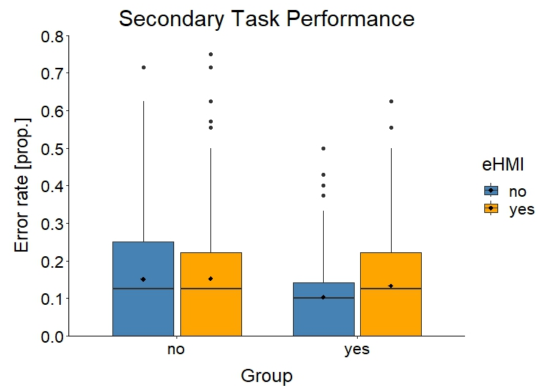

Figure 9: Error rate in the secondary task on the smartphone (n-back) by *eHMI* communication and *crossing group*. Smaller values represent fewer errors. More errors with *eHMI* than without. For the details of the significance test, see the text.

the AVs were equipped with external communication compared to when not, participants made more errors.

4.5 Perceived Criticality and Perceived Safety

Perceived criticality was measured on a 7-point Likert scale, ratings ranged from *Min* = 1 to *Max* = 7, with higher values representing higher perceived criticality. On average, participants perceived the *eHMI*-only condition as least critical (M = 1.40, SD = 0.65) and the *smartphone*-only condition as the most critical (M = 3.28, SD = 1.39; see Figure 10 left). Perceived safety was measured on a 5-point Likert scale, ratings ranged from *Min* = 1 to *Max* = 5, with higher values representing higher perceived safety. On average, participants felt safest when only the *eHMI* was active (M = 4.72, SD = 0.47) and least safe when they only had to do the distracting *smartphone* task (M = 3.54, SD = 0.99; see Figure 10 right).

The inferential analysis revealed a significant two-way interaction between *smartphone distraction* and *eHMI* for perceived criticality (β = 0.08, t = 2.49, p = .013) and perceived safety (β = -0.05, t = -2.53, p = .012). In addition, there was a significant two-way interaction between *smartphone distraction* and *crossing group* for perceived criticality (β = 0.07, t = 2.27, p = .024) and perceived safety (β = 0.06, t = 2.55, p = .011). When participants were **not distracted** by the secondary task on the smartphone, with *eHMI* compared to without *eHMI* led to lower perceived criticality (β = -0.20, t = -6.27, p < .001) and higher perceived safety (β = 0.16, t = 5.81, p < .001), while the *crossing group* had no influence on perceived criticality and perceived safety. When participants **were distracted by the smartphone** however, the *crossing group* led to lower perceived criticality (β = -0.14, t = -3.27, p = .001), and higher perceived safety (β = 0.11, t = 3.64, p < .001), while the *eHMI* had no impact on perceived criticality and perceived safety.

4.6 Perceived Workload

Perceived workload was subdivided into mental, physical, and temporal demand, performance, effort, and frustration. Ratings ranged from *Min* = 1 to *Max* = 21 for mental demand, temporal demand,

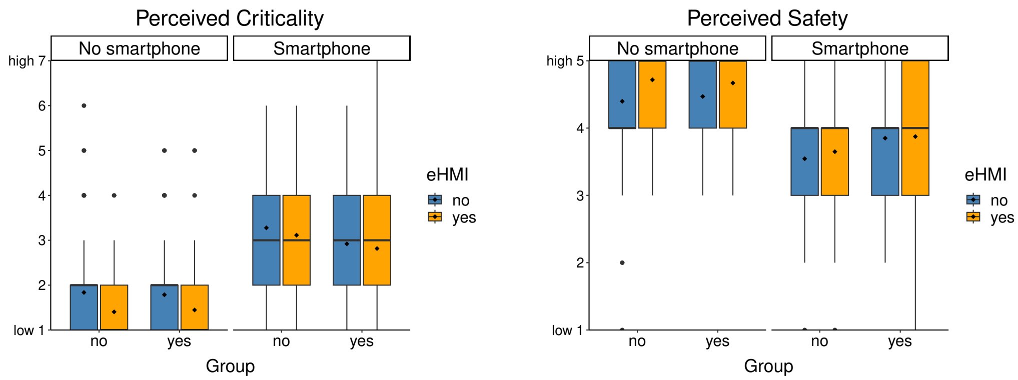

Figure 10: Perceived criticality (left) and perceived safety (right) by *eHMI* communication, *crossing group*, and *smartphone distraction*. Two-way interactions of *smartphone distraction* and *eHMI*, and *smartphone distraction* and *crossing group* for perceived criticality and perceived safety. For the details of the significance tests, see the text.

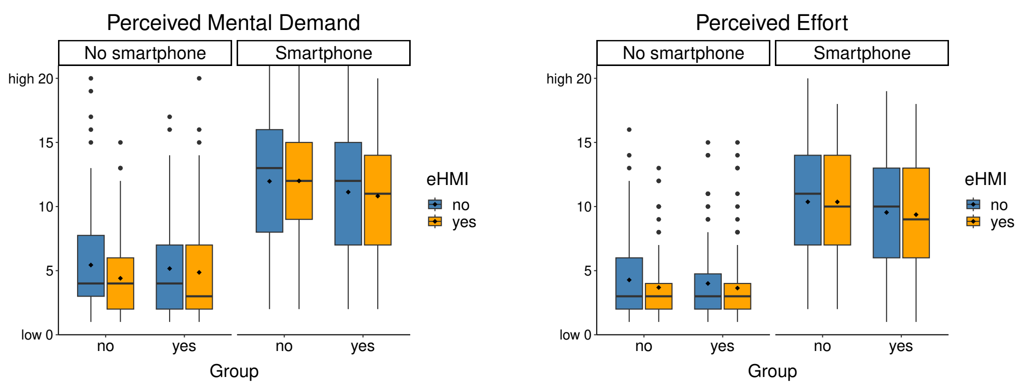

Figure 11: Perceived mental demand (left) and perceived effort (right) by *eHMI* communication, *crossing group*, and *smartphone distraction*. Two-way interaction of *crossing group* and *smartphone distraction* for mental demand and perceived effort. For the details of the significance tests, see the text.

performance and frustration, and from *Min* = 1 to *Max* = 20 for physical demand and effort. The results are grouped according to effect patterns, starting with mental demand and effort, followed by frustration, and lastly by physical demand, temporal demand, and performance.

4.6.1 *Mental Demand and Effort.*

In Figure 11 (left), the values for mental demand, and in Figure 11 (right), the values for perceived effort are depicted. There was a significant two-way interaction of *smartphone distraction* and the *group* crossing in front on mental demand (β = -0.27, t = -2.78, p = .006) and effort (β = -0.21, t = -2.23, p = .026). When the participants were **not distracted by a smartphone**, mental demand (β = -0.39, t = -3.72, p < .001) and effort (β = -0.27, t = -3.00, p = .003) were lower when the vehicles were equipped with *eHMIs* while the *crossing group* had no effect. When the participants were **distracted by a smartphone**, the *crossing group* led to lower mental demand (β = -0.45, t = -3.49, p < .001) and lower effort (β = -0.40, t = -3.12, p = .002) while the *eHMI* had no effect.

4.6.2 *Frustration.*

The first plot in the top left corner of Figure 12 displays the values for perceived frustration. There was a significant main effect of *smartphone distraction* (β = 1.79, t = 17.30, p < .001). The two-way interaction between *smartphone distraction* and *eHMI* (β = 0.18, t = 1.96, p = .050) and *smartphone distraction* and *group* did not reach significance (β = -0.18, t = -1.78, p = .078). On a descriptive level, when the participants were not distracted by a smartphone, the eHMI decreased the perceived frustration of participants. When participants were distracted by a smartphone, however, the crossing group decreased the perceived frustration.

4.6.3 *Perceived Performance, Physical and Temporal Demand.*

On a descriptive level, participants felt that their performance was overall good (see Figure 12.2), their physical demand was overall low (see Figure 12.3) and their temporal demand was medium to

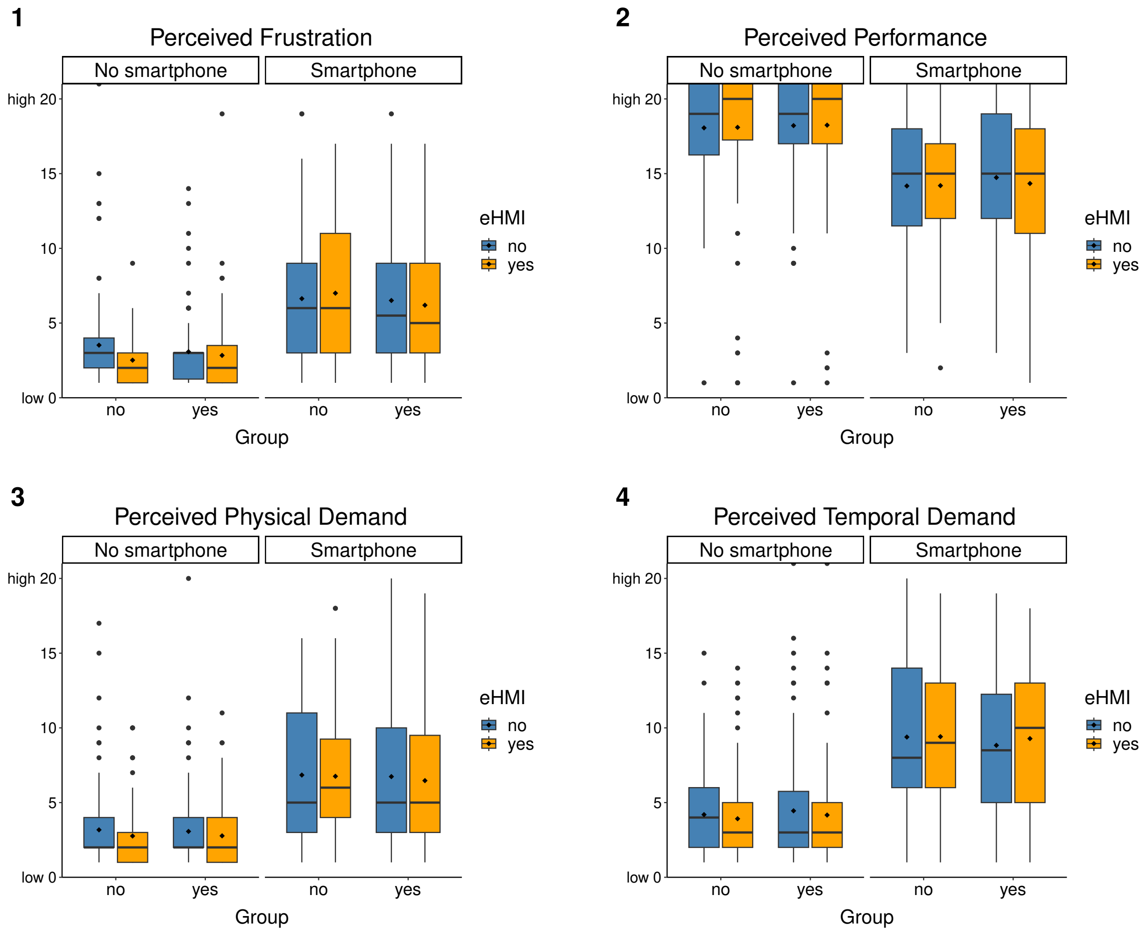

Figure 12: Perceived frustration (1), perceived performance (2), perceived physical demand (3) and perceived temporal demand (4) by *eHMI* communication, *crossing group* and *smartphone distraction*. Higher perceived frustration, physical demand, and temporal demand and lower perceived performance with *smartphone distraction*. For the details of the significance tests, see the text.

low (see Figure 12.4). Inferential analyses revealed no significant interaction effects for these variables. Only a significant main effect of *smartphone distraction* was found for the perceived physical demand ($\beta = 1.85$, $t = 23.25$, $p < .001$), the perceived temporal demand ($\beta = 2.58$, $t = 22.85$, $p < .001$) and the perceived performance ($\beta = -1.90$, $t = -17.15$, $p < .001$). The distraction task on the smartphone lead to higher physical and temporal demand and reduced perceived performance.

4.7 Interview

When participants were asked what influenced their crossing decision, 90.4% reported the *traffic* (e.g., gap size, vehicle speed, traffic flow), 56.5% reported the *group*, 23.5% reported the *eHMI* and 20.0% the *smartphone*. Numbers increased when the participants were explicitly asked about each of the three factors and whether it influenced their crossing decision. The highest increase had the *smartphone distraction* from 20.0% to 80.0%. *The group* and the *eHMI* had slightly lower numbers of 65.2% and 59.1% respectively. Participants said that when they were distracted by the *smartphone*, it was "more strenuous to make a decision" (P5), they were "more careful and looked more often" (P4) and "felt less safe" (P30). When asked about the influence of the *group*, participants stated that they "felt safer when others were around" (P28), "made a faster decision to cross" (P57) or stated that "the group was only helpful when I was using the smartphone"(P49). Regarding following the *group*, for example, participant 44 said that "when the group was crossing, I knew I could also cross". The *eHMI* was described by participants as "helpful" (P19) and that they "feel safer when the light was blue" (P48). While one participant (P34) said they "did not rely" on the *eHMI*, another participant said that they "were only crossing when the light was on" (P33). The majority of participants (91.3%) said that they were able to show their normal crossing behavior. The remaining ten participants said that they would normally not use a smartphone in traffic ($n = 5$), would cross in a smaller gap ($n = 4$), or would walk faster (n = 1). One participant said that they were tired, which led to a more cautious crossing behavior. Eleven participants recognized that the traffic or the gap was the same in all scenarios.

5 DISCUSSION

The main focus of this research was to examine how different factors that influence pedestrians' crossing perception and behavior interact with each other. This has been identified as a so far missing issue in the current body of research on pedestrians' interactions with AVs [16, 26, 50]. In this VR study, three factors were manipulated that already proved to be influential on their own in previous studies. The first factor was whether the AVs were equipped with *eHMIs* signalizing to the pedestrian if they would yield or not. The second factor was whether the pedestrian was distracted by doing a secondary task on a *smartphone* or not. And the third factor was whether there was another *group of pedestrians* that crossed in front of the participant or not. Several interaction effects were found in addition to main effects. When interaction effects were found, these are interpreted instead of the hypotheses regarding the main effects (applies to all hypotheses except for*H1a, H1b, H1c*; see Table 1).

Almost all interactions that were found included the factor *smartphone distraction*, either with external communication or with the crossing group. Therefore, attentiveness operationalized via smartphone distraction seems to be a key determinant in pedestrians' perception of and behavior during crossings.

5.1 Positive Effects of eHMI Disappear when Pedestrians are Distracted

When the pedestrians were attentive and not distracted by the smartphone, the eHMI had several positive effects such as a decrease in crossing duration, perceived criticality, mental demand, and effort and an increase in perceived safety. This is consistent with previous research with non-distracted pedestrians that also found these positive effects [8, 19, 31, 64], with the exception of crossing duration, where previous research found no effect of eHMIs [30]. When pedestrians were distracted by the smartphone, however, the positive effects of the eHMI disappeared, and no differences were found for these variables. Up to the authors' knowledge, this is the first study that demonstrates that the previously found positive main effects of the eHMIs when examined with attentive pedestrians do not apply for distracted pedestrians. This means that eHMIs seem to be less useful for pedestrian crossing behavior in more realistic scenarios. This raises the question of whether it makes sense to use eHMIs in such situations, whether new eHMI concepts are needed that are more beneficial in these circumstances, or whether implicit vehicle movements are sufficient.

Possible explanations why the positive effects of eHMIs disappeared when pedestrians were distracted could be that (1) they did not perceive the eHMI altogether or (2) had no capacity to process the provided information. The gaze data shows that distracted pedestrians did indeed look less often at traffic, but they looked proportionally longer at traffic when they were distracted and the eHMIs were active. This also provides a possible explanation why they made more errors in the smartphone task when the eHMIs were active. The eHMI drew more visual attention towards the vehicles as these are salient stimuli, so that participants spent less time on the smartphone and made more errors. This could be interpreted as either beneficial because the attention of pedestrians was focused more on traffic but it could also be disadvantageous as it adds more visual load [20]. Furthermore, the second possible explanation assumes that the eHMIs had no effect when pedestrians were distracted by the smartphone as they did not have enough capacity to correctly perceive, understand and interpret the additional eHMI information. Pedestrians indeed looked longer at traffic when the eHMIs were active and they were distracted by a smartphone which could be an indication that they took longer to process the additional information provided by the eHMI. However, this is not reflected in their subjective assessment of mental demand.

Overall, the pedestrians felt a higher physical and temporal demand, felt they performed worse and were more frustrated when they were distracted by the smartphone. They also looked less often and long at traffic. This is in line with previous research [43, 52, 85, 93].

5.2 Distracted Pedestrians Benefit from Crossing Group

As the interaction effects with the crossing group suggest, distracted pedestrians might rely on more familiar cues like other crossing pedestrians compared to the external communication of AVs when their cognitive resources are limited. The results show that the crossing group had several positive effects such as a higher perceived safety, lower perceived criticality, mental demand and effort when the pedestrians were distracted by the smartphone. No effects of the crossing group were found when the pedestrians were not distracted by the smartphone. This implies that the crossing group is not beneficial when pedestrians are attentive as they presumably then do not need the additional crossing information provided by the group. When pedestrians were distracted by the smartphone, however, they benefited from other pedestrians crossing in front. This is also reflected in pedestrians' gaze behavior as they looked less at the stopping AV when the group of other pedestrians was crossing in front. As Hamann et al. [40] hypothesized, following a group could be a way to compensate for the load that is induced by a distracting activity in a way that relying on the groups' social information is a mechanism to spare one's own resources.

5.3 Faster Crossing Initiation with eHMI or Pedestrian Group

As for the crossing initiation time, pedestrians initiated their crossing sooner when the eHMI was active (*H1a*) and when there was another group of pedestrians crossing (*H1b*). They initiated their crossing later when they were distracted by a smartphone (*H1c*). This supports the hypotheses and is consistent with previous findings on smartphone distraction, eHMIs and the presence of other pedestrians crossing [6, 8, 31, 33, 37, 47, 58, 71]. That pedestrians initiated their crossing sooner when other pedestrians were crossing could indicate the effect of responsibility diffusion [18]. As pedestrians followed others across the road more quickly, they might rely on them to accurately check traffic and make a safe crossing decision rather than doing it themselves. This can lead to dangerous situations if the other pedestrians misjudge the traffic situation, misunderstand the communication of the AV, or do not obey the traffic rules.

5.4 First Contact Effects for Subjective Variables

Lastly, first contact effects were found for the subjective variables. When first encountering a crossing scenario with AVs, the pedestrians felt less safe, perceived the situation as more critical, had a higher mental and temporal demand, higher effort and frustration than in later trials. Although there were practice trials, they were insufficient to avoid the effects mentioned. This underlines the importance of presenting the situation multiple times to give participants the chance to familiarize with the setting. A within-subjects design then seems to be particularly advantageous because no other time or learning effects were found for the crossing and gaze behavior. This is particularly noteworthy as some of the respondents recognized that they were in the same scenario, apart from the manipulated variables, but this did not alter their crossing behavior.

5.5 Limitations

We used a VR setting in this study, so transferability to a real-world setting is limited [83]. One potential issue with VR is that depths are commonly underestimated [29] which can affect the crossing decision. However, in this study, the participants only crossed as soon as the AV from the left came to a stop. Thus, the estimation of distances played a subordinate role. While some studies show that findings from pedestrian safety research in VR are comparable to real-world settings [22], it should be done with caution, and only relative and not absolute values should be used. As Holländer et al. [50] also point out, there is currently no ethical way to conduct field experiments where pedestrians are distracted by a smartphone while crossing the street. Thus, the VR setting seems to be a viable solution that is also widely used in eHMI research [8, 47, 71]. Even though we based our eHMI on previously studied eHMIs [8, 24, 31], results might differ with other eHMI designs or features such as an earlier onset of intention communication or a text-based eHMI. Nevertheless, we think this is a starting point for further research on how pedestrian factors influence the perception of eHMIs. In our study, we used a standardized distraction task instead of more realistic smartphone activities like texting or browsing. This allowed us to also examine the performance in the secondary task. Future studies should also investigate other secondary tasks as well as other modalities of the task (e.g., auditive [90]). Our sample was relatively young, so the generalizability of our results to other age groups is limited. Older people's ability to safely walk is even more affected by smartphone distraction [1, 78], so the effects could be more pronounced. Since we used a very standardized scenario, a few participants noticed that it was always the same gap in which they were crossing. This made it possible to keep the waiting time and the gap size the same, as these influence pedestrians' crossing decision [3, 97]. However, as the analyses show, this knowledge did not alter the participants' crossing and gaze behavior except for improving in the secondary task on the smartphone as well as their perceived performance. Finally, we have only considered a limited number of influencing factors. We have chosen these because they influence the uptake of information in the decision-making process and, apart from eHMIs, are frequently observed in reality. However, there are several other influencing factors, such as traffic and pedestrian density, or visibility of other road users, that play a role and should be investigated in future studies.

5.6 Implications and Future Work

The results of this study show that it is necessary to examine eHMI concepts in more complex and realistic crossing scenarios that consider pedestrian factors such as smartphone distraction or other pedestrians' behavior. As traffic is a complex system where many factors influence each other, interaction effects can be expected. It was a logical first step to investigate a new technology like eHMIs in simplified scenarios first. However, those results cannot necessarily be transferred to other, more complex contexts where there is not only one AV and one attentive pedestrian. As others have pointed out [16, 26], the next step is to apply these findings to more realistic crossing scenarios that also consider pedestrian factors as this study has done. The results affirm that previously found main effects cannot simply be transferred to more complex situations and that it is necessary to examine interaction effects. The study shows that eHMIs help in idealised pedestrian situations but are less helpful in more complex traffic situations and in particular when the pedestrian is distracted by a smartphone. Accordingly, the attentional state of the pedestrian is a key factor that should be further investigated.. More research is needed to better understand the extent to which visual or cognitive workload is responsible for the effects.

For pedestrians who are distracted by their smartphone, the position of the eHMI on the vehicle might not be ideal. Other solutions, such as on the sidewalk or on the smartphone [49, 50, 69, 79] might be more promising. However, research shows that pedestrians ability to detect in-ground signals is rather late, so this approach might also not be particularly useful [55]. Using the smartphone, on the other hand, could lead to even more distracted walking and reduced situational awareness since people rely on the technology too much [79]. In addition, these solutions do not account for other types of distraction such as talking to others, daydreaming, or eating.

Another point to examine further is whether pedestrians rely more on social cues than on technical cues when their resources are diminished by another task. This could be done by providing conflicting information from the social and technical cues as Colley et al. [8] have done, to see what information pedestrians rely on. More studies are needed to better understand if this is due to a higher familiarity with social cues or not enough trust in the technical cues.

Overall, it is necessary to continue the discussion under which circumstances and in which situations eHMIs are needed and beneficial in contrast to already existing implicit cues like vehicle movement patterns. Currently, the majority of studies find positive effects of eHMIs in rather idealized scenarios, but there are also studies that already point to potential problems of eHMIs in real traffic (e.g., scalability issues). However, there is a lack of empirical research on the circumstances under which eHMIs are necessary and helpful. More research is needed to better target the use of eHMIs in situations where they are beneficial. The results of this study provide initial insights into how the influencing factors of smartphone distraction and the presence of other pedestrians affect crossing and gaze behavior in automated traffic. Furthermore, the results of this study can be used to develop new concepts for pedestrian-AV interaction, as the existing concepts do not seem to be universally applicable.

5.7 The Way Forward

The inclusion of pedestrian factors in eHMI research is only just beginning. This study provides evidence of how important it is to include pedestrian attention in particular, as this factor was involved in almost all interactions that were found. In the future, the distraction of pedestrians, not only by smartphones but also by other sources, should be looked at more closely as it is a safety critical issue with increasing accident numbers [75, 82]. As Rasouli and Tsotsos [81] point out, there are 38 factors that influence pedestrians' decision-making process at the point of crossing. While all factors have an influence, some should be given priority and investigated in upcoming studies. As others have already pointed out [16, 26] and as we implemented in our study, the presence of other road users is important. Traffic is a system with many actors whose behavior influences each other. This means not only including more than one pedestrian or more than one vehicle but also bicyclists, motorcyclists, and public transport vehicles. The relation between those actors also matters, as, for example, standing next to a group of strangers or being with a group of people you know changes your behavior in traffic [50, 61]. As some road users are especially vulnerable, such as children, the elderly, and people with impairments [39, 48], they should not be left out of the research and development of automated driving in cities. Furthermore, it is necessary to find an appropriate degree of standardization for external communication of AVs [54]. For this, we need realistic and relevant scenarios and factors so that the results also hold up in the real world outside the laboratory or test tracks.

6 CONCLUSION

This work investigated how the existing effects of eHMIs transfer to a more realistic setting where the pedestrian is distracted by a smartphone and where other pedestrians also cross. For this, we conducted a VR study with 115 participants to explore interaction effects between intent-based *eHMIs* of AVs, pedestrians' *smartphone distraction* and the presence of *other crossing pedestrians*. Interaction effects were found especially in regard to pedestrians' smartphone distraction. We could replicate several positive effects that eHMIs have on attentive pedestrians. When AVs communicated their intent via an eHMI, attentive pedestrians crossed faster and perceived the situation to be less critical, felt safer, and had a lower cognitive workload and perceived effort. However, these effects disappeared when pedestrians were distracted by a smartphone. Rather than relying on the technical cues provided by the eHMI, distracted pedestrians relied on the social cues provided by the group crossing the street. They reported a lower cognitive workload and effort, felt safer, and perceived the situation as less critical when there was another group of pedestrians crossing. The pedestrians also looked less often and less long at the stopping AV, which communicated its intent with the eHMI when a group was present. In terms of performance on the secondary task, pedestrians made more errors when the eHMIs were active, which can be related to the fact that they looked proportionally longer at traffic when they were distracted and the eHMIs were active. In addition, they initiated the crossing faster when other pedestrians were crossing in front or the eHMIs were active and slower when they were using a smartphone. Our results provide evidence that pedestrian factors interact with eHMI effects and that more research is needed on complex, more realistic pedestrian-AV crossing scenarios.

ACKNOWLEDGMENTS

We would like to thank Paulina Kolland and Lukas Braun for their support with data collection, as well as Jens Spannagel, Tobias Wagner, Jan Henry Belz, and Tim Fabian for their support in the implementation. We would also like to thank all study participants.

REFERENCES

[1] Linson J. Alapatt, Nancye M. Peel, Natasha Reid, Leonard C. Gray, and Ruth E. Hubbard. 2020. The Effect of Age on Gait Speed When Texting. *International Journal of Environmental Research and Public Health* 17, 2 (2020), 599. https://doi.org/10.3390/ijerph17020599

[2] Rushdi Alsaleh, Tarek Sayed, and Mohamed H. Zaki. 2018. Assessing the Effect of Pedestrians' Use of Cell Phones on Their Walking Behavior: A Study Based on Automated Video Analysis. *Transportation Research Record* 2672, 35 (2018), 46–57. https://doi.org/10.1177/0361198118780708

[3] Harley Amado, Sara Ferreira, José Pedro Tavares, Paulo Ribeiro, and Elisabete Freitas. 2020. Pedestrian–Vehicle Interaction at Unsignalized Crosswalks: A Systematic Review. *Sustainability* 12, 7 (2020), 2805. https://doi.org/10.3390/su12072805

[4] Ashratuz Zavin Asha, Christopher Smith, Georgina Freeman, Sean Crump, Sowmya Somanath, Lora Oehlberg, and Ehud Sharlin. 2021. Co-Designing Interactions between Pedestrians in Wheelchairs and Autonomous Vehicles. In *Designing Interactive Systems Conference 2021* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 339–351. https://doi.org/10.1145/3461778.3462068

[5] P. Bazilinskyy, L. Kooijman, D. Dodou, and J.C.F. de Winter. 2021. How should external human-machine interfaces behave? Examining the effects of colour, position, message, activation distance, vehicle yielding, and visual distraction among 1,434 participants. *Applied Ergonomics* 95 (2021), 103450. https://doi.org/10.1016/j.apergo.2021.103450

[6] Katherine W. Byington and David C. Schwebel. 2013. Effects of mobile Internet use on college student pedestrian injury risk. *Accident Analysis & Prevention* 51 (2013), 78–83. https://doi.org/10.1016/j.aap.2012.11.001

[7] Siyuan Chen, Julien Epps, Natalie Ruiz, and Fang Chen. 2011. Eye Activity as a Measure of Human Mental Effort in HCI. In *Proceedings of the 16th International Conference on Intelligent User Interfaces* (Palo Alto, CA, USA) *(IUI '11)*. Association for Computing Machinery, New York, NY, USA, 315–318. https://doi.org/10.1145/1943403.1943454

[8] Mark Colley, Elvedin Bajrovic, and Enrico Rukzio. 2022. Effects of Pedestrian Behavior, Time Pressure, and Repeated Exposure on Crossing Decisions in Front of Automated Vehicles Equipped with External Communication. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 367, 11 pages. https://doi.org/10.1145/3491102.3517571

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

[10] Mark Colley, Christian Hummler, and Enrico Rukzio. 2022. Effects of mode distinction, user visibility, and vehicle appearance on mode confusion when interacting with highly automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour* 89 (2022), 303–316. https://doi.org/10.1016/j.trf.2022.06.020

[11] Mark Colley, Surong Li, and Enrico Rukzio. 2021. Increasing Pedestrian Safety Using External Communication of Autonomous Vehicles for Signalling Hazards. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction* (Toulouse & Virtual, France) *(MobileHCI '21)*. Association for Computing Machinery, New York, NY, USA, Article 20, 10 pages. https://doi.org/10.1145/3447526.3472024

[12] Mark Colley, Stefanos Can Mytilineos, Marcel Walch, Jan Gugenheimer, and Enrico Rukzio. 2020. Evaluating Highly Automated Trucks as Signaling Lights. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '20)*. Association for Computing Machinery, New York, NY, USA, 111–121. https://doi.org/10.1145/3409120.3410647

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

[14] Mark Colley, Marcel Walch, Jan Gugenheimer, Ali Askari, and Enrico Rukzio. 2020. Towards Inclusive External Communication of Autonomous Vehicles for Pedestrians with Vision Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376472

[15] Mark Colley, Marcel Walch, and Enrico Rukzio. 2019. For a Better (Simulated) World: Considerations for VR in External Communication Research. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 442–449. https://doi.org/10.1145/3349263.3351523

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

[17] Mark Colley, Bastian Wankmüller, Tim Mend, Thomas Väth, Enrico Rukzio, and Jan Gugenheimer. 2022. User gesticulation inside an automated vehicle with external communication can cause confusion in pedestrians and a lower willingness to cross. *Transportation Research Part F: Traffic Psychology and Behaviour* 87 (2022), 120–137. https://doi.org/10.1016/j.trf.2022.03.011

[18] John M. Darley and Bibb Latane. 1968. Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology* 8, 4 (1968), 377–383. https://doi.org/10.1037/h0025589

[19] Koen de Clercq, Andre Dietrich, Juan Pablo Núñez Velasco, Joost de Winter, and Riender Happee. 2019. External Human-Machine Interfaces on Automated Vehicles: Effects on Pedestrian Crossing Decisions. *Human Factors* 61, 8 (2019), 1353–1370. https://doi.org/10.1177/0018720819836343

[20] Joost de Winter and Dimitra Dodou. 2022. External human–machine interfaces: Gimmick or necessity? *Transportation Research Interdisciplinary Perspectives* 15 (2022), 100643. https://doi.org/10.1016/j.trip.2022.100643

[21] Shuchisnigdha Deb, Daniel W. Carruth, Muztaba Fuad, Laura M. Stanley, and Darren Frey. 2020. Comparison of Child and Adult Pedestrian Perspectives of External Features on Autonomous Vehicles Using Virtual Reality Experiment. In *Advances in Human Factors of Transportation*, Neville Stanton (Ed.). Springer International Publishing, Cham, 145–156. https://doi.org/10.1007/978-3-030-20503-4_13

[22] Shuchisnigdha Deb, Daniel W. Carruth, Richard Sween, Lesley Strawderman, and Teena M. Garrison. 2017. Efficacy of virtual reality in pedestrian safety research. *Applied Ergonomics* 65 (2017), 449–460. https://doi.org/10.1016/j.apergo.2017.03.007

[23] Debargha Dey, Azra Habibovic, Andreas Löcken, Philipp Wintersberger, Bastian Pfleging, Andreas Riener, Marieke Martens, and Jacques Terken. 2020. Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transportation Research Interdisciplinary Perspectives* 7 (2020), 100174. https://doi.org/10.1016/j.trip.2020.100174

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

[25] Debargha Dey, Andrii Matviienko, Melanie Berger, Bastian Pfleging, Marieke Martens, and Jacques Terken. 2021. Communicating the intention of an automated vehicle to pedestrians: The contributions of eHMI and vehicle behavior. *it - Information Technology* 63, 2 (2021), 123–141. https://doi.org/10.1515/itit-2020-0025

[26] Debargha Dey, Arjen van Vastenhoven, Raymond H. Cuijpers, Marieke Martens, and Bastian Pfleging. 2021. Towards Scalable EHMIs: Designing for AV-VRU Communication Beyond One Pedestrian. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) *(AutomotiveUI '21)*. Association for Computing Machinery, New York, NY, USA, 274–286. https://doi.org/10.1145/3409118.3475129

[27] Joshua E. Domeyer, John D. Lee, and Heishiro Toyoda. 2020. Vehicle Automation–Other Road User Communication and Coordination: Theory and Mechanisms. *IEEE Access* 8 (2020), 19860–19872. https://doi.org/10.1109/ACCESS.2020.2969233

[28] Aurélie Dommes, M-A Granié, M-S Cloutier, Cécile Coquelet, and Florence Huguenin-Richard. 2015. Red light violations by adult pedestrians and other safety-related behaviors at signalized crosswalks. *Accident Analysis & Prevention* 80 (2015), 67–75. https://doi.org/10.1016/j.aap.2015.04.002

[29] Fatima El Jamiy and Ronald Marsh. 2019. Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality. *IET Image Processing* 13, 5 (2019), 707–712. https://doi.org/10.1049/iet-ipr.2018.5920

[30] Stefanie M Faas and Martin Baumann. 2019. Yielding light signal evaluation for self-driving vehicle and pedestrian interaction. In *Human Systems Engineering and Design II. IHSED 2019. Advances in Intelligent Systems and Computing*, Vol. 1026. Springer, 189–194. https://doi.org/10.1007/978-3-030-27928-8_29

[31] Stefanie M. Faas, Andrea C. Kao, and Martin Baumann. 2020. A Longitudinal Video Study on Communicating Status and Intent for Self-Driving Vehicle Pedestrian Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376484

[32] Daniel J Fagnant and Kara Kockelman. 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice* 77 (2015), 167–181. https://doi.org/10.1016/j.tra.2015.04.003]

[33] Jolyon J. Faria, Stefan Krause, and Jens Krause. 2010. Collective behavior in road crossing pedestrians: the role of social information. *Behavioral Ecology* 21, 6 (09 2010), 1236–1242. https://doi.org/10.1093/beheco/arq141

[34] Nicholas N Ferenchak and Matin Katirai. 2017. Pedestrian Crossing Behavior in Relation to Grouping and Gender in a Developing Country Context. *Journal of Global Epidemiology and Environmental Health 2017* (2017), 37–45. https://doi.org/10.29199/2637-7144/GEEH-101018

[35] Cesar Fernandez, Maria A. Vicente, Irene Carrillo, Mercedes Guilabert, and José J. Mira. 2020. Factors influencing the smartphone usage behavior of pedestrians: Observational study on "Spanish Smombies". *Journal of Medical Internet Research* 22, 8 (2020), e19350. https://doi.org/10.2196/19350

[36] Duane R Geruschat, Shirin E Hassan, and Kathleen A Turano. 2003. Gaze behavior while crossing complex intersections. *Optometry and vision science* 80, 7 (2003), 515–528.

[37] George Gillette, Kay Fitzpatrick, Susan Chrysler, and Raul Avelar. 2016. Effect of Distractions on a Pedestrian's Waiting Behavior at Traffic Signals: Observational Study. *Transportation Research Record* 2586, 1 (2016), 111–119. https://doi.org/10.3141/2586-13

[38] Marie-Axelle Granié, Marjorie Pannetier, and Ludivine Gueho. 2013. Developing a self-reporting method to measure pedestrian behaviors at all ages. *Accident Analysis & Prevention* 50 (2013), 830–839. https://doi.org/10.1016/j.aap.2012.07.009

[39] Mathias Haimerl, Mark Colley, and Andreas Riener. 2022. Evaluation of Common External Communication Concepts of Automated Vehicles for People With Intellectual Disabilities. *Proc. ACM Hum.-Comput. Interact.* 6, MHCI, Article 182 (sep 2022), 19 pages. https://doi.org/10.1145/3546717

[40] Cara Hamann, Diana Dulf, Erika Baragan-Andrada, Morgan Price, and Corinne Peek-Asa. 2017. Contributors to pedestrian distraction and risky behaviours during road crossings in Romania. 23, 6 (2017), 370–376. https://doi.org/10.1136/injuryprev-2016-042219

[41] W. Andrew Harrell. 1991. Factors influencing pedestrian cautiousness in crossing streets. *Journal of Social Psychology* 131, 3 (1991), 367–372. https://doi.org/10.1080/00224545.1991.9713863

[42] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[43] Julie Hatfield and Susanne Murphy. 2007. The effects of mobile phone use on pedestrian crossing behaviour at signalised and unsignalised intersections. *Accident Analysis & Prevention* 39, 1 (2007), 197–205. https://doi.org/10.1016/j.aap.2006.07.001

[44] Samantha H. Haus, Rini Sherony, and Hampton C. Gabler. 2019. Estimated benefit of automated emergency braking systems for vehicle–pedestrian crashes in the United States. *Traffic Injury Prevention* 20, Sup1 (2019), S171–S176. https://doi.org/10.1080/15389588.2019.1602729

[45] Ann-Christin Hensch, Isabel Kreißig, Matthias Beggiato, and Josef F. Krems. 2022. The Effect of eHMI Malfunctions on Younger and Elderly Pedestrians' Trust and Acceptance of Automated Vehicle Communication Signals. *Frontiers in Psychology* 13 (2022), 866475. https://doi.org/10.3389/fpsyg.2022.866475

[46] Ann-Christin Hensch, Isabel Neumann, Matthias Beggiato, Josephine Halama, and Josef F Krems. 2019. Effects of a light-based communication approach as an external HMI for Automated Vehicles–a Wizard-of-Oz Study. *Transactions on Transport Sciences* 10, 2 (2019), 18–32. https://doi.org/10.5507/tots.2019.012

[47] Kai Holländer, Ashley Colley, Christian Mai, Jonna Häkkilä, Florian Alt, and Bastian Pfleging. 2019. Investigating the Influence of External Car Displays

on Pedestrians' Crossing Behavior in Virtual Reality. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (Taipei, Taiwan) *(MobileHCI '19)*. Association for Computing Machinery, New York, NY, USA, Article 27, 11 pages. https://doi.org/10.1145/3338286.3340138

[48] Kai Holländer, Mark Colley, Enrico Rukzio, and Andreas Butz. 2021. A Taxonomy of Vulnerable Road Users for HCI Based On A Systematic Literature Review. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 158, 13 pages. https://doi.org/10.1145/3411764.3445480

[49] Kai Holländer, Marius Hoggenmüller, Romy Gruber, Sarah Theres Völkel, and Andreas Butz. 2022. Take It to the Curb: Scalable Communication Between Autonomous Cars and Vulnerable Road Users Through Curbstone Displays. *Frontiers in Computer Science* 4 (2022). https://doi.org/10.3389/fcomp.2022.844245

[50] Kai Holländer, Andy Krüger, and Andreas Butz. 2020. Save the Smombies: App-Assisted Street Crossing. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services* (Oldenburg, Germany) *(MobileHCI '20)*. Association for Computing Machinery, New York, NY, USA, Article 22, 11 pages. https://doi.org/10.1145/3379503.3403547

[51] Tim Horberry, Rachel Osborne, and Kristie Young. 2019. Pedestrian smartphone distraction: prevalence and potential severity. *Transportation Research Part F: Traffic Psychology and Behaviour* 60 (2019), 515–523. https://doi.org/10.1016/j.trf.2018.11.011

[52] Kang Jiang, Feiyang Ling, Zhongxiang Feng, Changxi Ma, Wesley Kumfer, Chen Shao, and Kun Wang. 2018. Effects of mobile phone distraction on pedestrians' crossing behavior and visual attention allocation at a signalized intersection: An outdoor experimental study. *Accident Analysis & Prevention* 115 (2018), 170–177. https://doi.org/10.1016/j.aap.2018.03.019

[53] Philip Joisten, Ziyu Liu, Nina Theobald, Andreas Webler, and Bettina Abendroth. 2021. Communication of Automated Vehicles and Pedestrian Groups: An Intercultural Study on Pedestrians' Street Crossing Decisions. In *Mensch Und Computer 2021* (Ingolstadt, Germany) *(MuC '21)*. Association for Computing Machinery, New York, NY, USA, 49–53. https://doi.org/10.1145/3473856.3474004

[54] Christina Kaß, Stefanie Schoch, Frederik Naujoks, Sebastian Hergeth, Andreas Keinath, and Alexandra Neukum. 2020. Standardized Test Procedure for External Human–Machine Interfaces of Automated Vehicles. *Information* 11, 3 (2020), 173. https://doi.org/10.3390/info11030173

[55] Eunjee Kim, Hyorim Kim, Yujin Kwon, Seobin Choi, and Gwanseob Shin. 2021. Performance of ground-level signal detection when using a phone while walking. *Accident Analysis & Prevention* 151 (2021), 105909. https://doi.org/10.1016/j.aap.2020.105909

[56] Hye-Jin Kim, Jin-Young Min, Hyun-Jin Kim, and Kyoung-Bok Min. 2017. Accident risk associated with smartphone addiction: A study on university students in Korea. *Journal of Behavioral Addictions* 6, 4 (2017), 699–707. https://doi.org/10.1556/2006.6.2017.070

[57] Julian FP Kooij, Fabian Flohr, Ewoud AI Pool, and Dariu M Gavrila. 2019. Context-based path prediction for targets with switching dynamics. *International Journal of Computer Vision* 127, 3 (2019), 239–262. https://doi.org/10.1007/s11263-018-1104-4

[58] Lars Kooijman, Riender Happee, and Joost C. F. de Winter. 2019. How Do eHMIs Affect Pedestrians' Crossing Behavior? A Study Using a Head-Mounted Display Combined with a Motion Suit. *Information* 10, 12 (2019), 386. https://doi.org/10.3390/info10120386

[59] Oswald D. Kothgassner, Anna Felnhofer, Nathalie Hauk, Elisabeth Kastenhofer, Jasmine Gomm, and Ilse Kryspin-Exner. 2012. *Technology Usage Inventory - Manual.* Retrieved September 15, 2022 from https://www.ffg.at/sites/default/files/allgemeine_downloads/thematische%20programme/programmdokumente/tui_manual.pdf

[60] Mirjam Lanzer, Franziska Babel, Fei Yan, Bihan Zhang, Fang You, Jianmin Wang, and Martin Baumann. 2020. Designing Communication Strategies of Autonomous Vehicles with Pedestrians: An Intercultural Study. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) *(AutomotiveUI '20)*. Association for Computing Machinery, New York, NY, USA, 122–131. https://doi.org/10.1145/3409120.3410653

[61] Mirjam Lanzer and Martin Baumann. 2020. Does crossing the road in a group influence pedestrians' gaze behavior? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64, 1 (2020), 1938–1942. https://doi.org/10.1177/1071181320641467

[62] Mirjam Lanzer, Tanja Stoll, Mark Colley, and Martin Baumann. 2021. Intelligent Mobility in the City: The Influence of System and Context Factors on Drivers' Takeover Willingness and Trust in Automated Vehicles. *Frontiers in Human Dynamics* 3 (2021). https://doi.org/10.3389/fhumd.2021.676667

[63] Grégoire S. Larue, Christopher N. Watling, Alexander A. Black, Joanne M. Wood, and Mahrokh Khakzar. 2020. Pedestrians distracted by their smartphone: Are in-ground flashing lights catching their attention? A laboratory study. *Accident Analysis & Prevention* 134 (2020), 105346. https://doi.org/10.1016/j.aap.2019.105346

[64] Merle Lau, Duc Hai Le, and Michael Oehl. 2021. Design of External Human-Machine Interfaces for Different Automated Vehicle Types for the Interaction with Pedestrians on a Shared Space. In *Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021)*, Nancy L. Black, W. Patrick Neumann, and Ian Noy (Eds.). Springer International Publishing, Cham, 710–717. https://doi.org/10.1007/978-3-030-74608-7_87

[65] Yee Mun Lee, Ruth Madigan, Oscar Giles, Laura Garach-Morcillo, Gustav Markkula, Charles Fox, Fanta Camara, Markus Rothmueller, Signe Alexandra Vendelbo-Larsen, Pernille Holm Rasmussen, Andre Dietrich, Dimitris Nathanael, Villy Portouli, Anna Schieben, and Natasha Merat. 2021. Road users rarely use explicit communication when interacting in today's traffic: implications for automated vehicles. *Cognition, Technology & Work* 23 (2021), 367–380. https://doi.org/10.1007/s10111-020-00635-y

[66] Sammy Licence, Robynne Smith, Miranda P. McGuigan, and Conrad P. Earnest. 2015. Gait Pattern Alterations during Walking, Texting and Walking and Texting during Cognitively Distractive Tasks while Negotiating Common Pedestrian Obstacles. *PLOS ONE* 10, 7 (07 2015), 1–11. https://doi.org/10.1371/journal.pone.0133281

[67] Ming-I Brandon Lin and Yu-Ping Huang. 2017. The impact of walking while using a smartphone on pedestrians' awareness of roadside events. *Accident Analysis & Prevention* 101 (2017), 87–96. https://doi.org/10.1016/j.aap.2017.02.005

[68] Hailong Liu, Takatsugu Hirayama, Luis Yoichi Morales Saiki, and Hiroshi Murase. 2022. Implicit Interaction with an Autonomous Personal Mobility Vehicle: Relations of Pedestrians' Gaze Behavior with Situation Awareness and Perceived Risks. *International Journal of Human–Computer Interaction* 0, 0 (2022), 1–17. https://doi.org/10.1080/10447318.2022.2073006

[69] Andreas Löcken, Carmen Golling, and Andreas Riener. 2019. How Should Automated Vehicles Interact with Pedestrians? A Comparative Analysis of Interaction Concepts in Virtual Reality. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 262–274. https://doi.org/10.1145/3342197.3344544

[70] Stefanie M. Faas, Johannes Kraus, Alexander Schoenhals, and Martin Baumann. 2021. Calibrating Pedestrians' Trust in Automated Vehicles: Does an Intent Display in an External HMI Support Trust Calibration and Safe Crossing Behavior?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 157, 17 pages. https://doi.org/10.1145/3411764.3445738

[71] Karthik Mahadevan, Elaheh Sanoubari, Sowmya Somanath, James E. Young, and Ehud Sharlin. 2019. AV-Pedestrian Interaction Design Using a Pedestrian Mixed Traffic Simulator. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19)*. Association for Computing Machinery, New York, NY, USA, 475–486. https://doi.org/10.1145/3322276.3322328

[72] Dylan Moore, Rebecca Currano, G. Ella Strack, and David Sirkin. 2019. The Case for Implicit External Human-Machine Interfaces for Autonomous Vehicles. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 295–307. https://doi.org/10.1145/3342197.3345320

[73] Mehdi Moussaïd, Niriaska Perozo, Simon Garnier, Dirk Helbing, and Guy Theraulaz. 2010. The Walking Behaviour of Pedestrian Social Groups and Its Impact on Crowd Dynamics. *PLoS ONE* 5, 4 (2010), e10047. https://doi.org/10.1371/journal.pone.0010047

[74] Judith Mwakalonge, Saidi Siuhi, and Jamario White. 2015. Distracted walking: Examining the extent to pedestrian safety problems. *Journal of Traffic and Transportation Engineering (English Edition)* 2, 5 (2015), 327–337. https://doi.org/10.1016/j.jtte.2015.08.004

[75] Jack L. Nasar and Derek Troyer. 2013. Pedestrian injuries due to mobile phone use in public places. *Accident Analysis & Prevention* 57 (2013), 91–95. https://doi.org/10.1016/j.aap.2013.03.021

[76] Dalibor Pešić, Boris Antić, Draženko Glavić, and Marina Milenković. 2016. The effects of mobile phone use on pedestrian crossing behaviour at unsignalized intersections–Models for predicting unsafe pedestrians behaviour. *Safety Science* 82 (2016), 1–8. https://doi.org/10.1016/j.ssci.2015.08.016

[77] Andrew J. Piazza, Adam P. Knowlden, Elizabeth Hibberd, James Leeper, Angelia M. Paschal, and Stuart Usdan. 2020. Distracted mobile device use among street-crossing college student pedestrians: An observational approach. *Journal of American College Health* 70, 7 (2020), 2135–2142. https://doi.org/10.1080/07448481.2020.1845182

[78] Paphawee Prupetkaew, Vipul Lugade, Teerawat Kamnardsiri, and Patima Silsupadol. 2019. Cognitive and visual demands, but not gross motor demand, of concurrent smartphone use affect laboratory and free-living gait among young and older adults. *Gait & Posture* 68 (2019), 30–36. https://doi.org/10.1016/j.gaitpost.2018.11.003

[79] Pooya Rahimian, Elizabeth E. O'Neal, Shiwen Zhou, Jodie M. Plumert, and Joseph K. Kearney. 2018. Harnessing Vehicle-to-Pedestrian (V2P) Communication Technology: Sending Traffic Warnings to Texting Pedestrians. *Human Factors* 60, 6 (2018), 833–843. https://doi.org/10.1177/0018720818781365

[80] Rahul Raoniar and Akhilesh Kumar Maurya. 2022. Pedestrian red-light violation at signalised intersection crosswalks: Influence of social and non-social factors. *Safety Science* 147 (2022), 105583. https://doi.org/10.1016/j.ssci.2021.105583

[81] Amir Rasouli and John K. Tsotsos. 2020. Autonomous Vehicles That Interact With Pedestrians: A Survey of Theory and Practice. *IEEE Transactions on Intelligent Transportation Systems* 21, 3 (2020), 900–918. https://doi.org/10.1109/TITS.2019.2901817

[82] Jun Ren, Yue Chen, Fenfen Li, Cheng Xue, Xiaoya Yin, Juanjuan Peng, Ji Liang, Qiming Feng, and Shumei Wang. 2021. Road Injuries Associated With Cellular Phone Use While Walking or Riding a Bicycle or an Electric Bicycle: A Case-Crossover Study. *American Journal of Epidemiology* 190, 1 (2021), 37–43. https://doi.org/10.1093/aje/kwaa164

[83] Sonja Schneider, Philipp Maruhn, Nguyen-Thong Dang, Prashant Pala, Viola Cavallo, and Klaus Bengler. 2021. Pedestrian Crossing Decisions in Virtual Environments: Behavioral Validity in CAVEs and Head-Mounted Displays. *Human Factors* 64, 7 (2021), 0018720820987446. https://doi.org/10.1177/0018720820987446

[84] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (06 2001), 266–281. https://doi.org/10.1162/105474601300343603

[85] Sarah M Simmons, Jeff K Caird, Alicia Ta, Franci Sterzer, and Brent E Hagel. 2020. Plight of the distracted pedestrian: a research synthesis and meta-analysis of mobile phone use on crossing behaviour. *Injury Prevention* 26, 2 (2020), 170–176. https://doi.org/10.1136/injuryprev-2019-043426

[86] Christos Stogios, Dena Kasraian, Matthew J. Roorda, and Marianne Hatzopoulou. 2019. Simulating impacts of automated driving behavior and traffic conditions on vehicle emissions. *Transportation Research Part D: Transport and Environment* 76 (2019), 176–192. https://doi.org/10.1016/j.trd.2019.09.020

[87] David L. Strayer, Jonna Turrill, Joel M. Cooper, James R. Coleman, Nathan Medeiros-Ward, and Francesco Biondi. 2015. Assessing Cognitive Distraction in the Automobile. *Human Factors* 57, 8 (2015), 1300–1324. https://doi.org/10.1177/0018720815575149

[88] Matus Sucha, Daniel Dostal, and Ralf Risser. 2017. Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention* 102 (2017), 41–50. https://doi.org/10.1016/j.aap.2017.02.018

[89] Leah L. Thompson, Frederick P. Rivara, Rajiv C. Ayyagari, and Beth E. Ebel. 2013. Impact of social and technological distraction on pedestrian crossing behaviour: An observational study. *Injury Prevention* 19, 4 (2013), 232–237. https://doi.org/10.1136/injuryprev-2012-040601

[90] Kai Tian, Gustav Markkula, Chongfeng Wei, Ehsan Sadraei, Toshiya Hirose, Natasha Merat, and Richard Romano. 2022. Impacts of visual and cognitive distractions and time pressure on pedestrian crossing behaviour: A simulator study. *Accident Analysis & Prevention* 174 (2022), 106770. https://doi.org/10.1016/j.aap.2022.106770

[91] Tom Tullis and Bill Albert. 2013. Chapter 7 - Behavioral and Physiological Metrics. In *Measuring the User Experience* (2 ed.), Tom Tullis and Bill Albert (Eds.). Morgan Kaufmann, Boston, 163–186. https://doi.org/10.1016/B978-0-12-415781-1.00007-8

[92] Roni Utriainen. 2021. The potential impacts of automated vehicles on pedestrian safety in a four-season country. *Journal of Intelligent Transportation Systems* 25, 2 (2021), 188–196. https://doi.org/10.1080/15472450.2020.1845671

[93] Mark Vollrath, Anja Katharina Huemer, and Christin Nicolai. 2019. Young people use their smartphone all the time–even when crossing the street? *IET Intelligent Transport Systems* 13, 8 (2019), 1213–1217. https://doi.org/10.1049/iet-its.2018.5548IET

[94] Nikolai von Janczewski, Jennifer Wittmann, Arnd Engeln, Martin Baumann, and Lutz Krauß. 2021. A meta-analysis of the n-back task while driving and its effects on cognitive workload. *Transportation Research Part F: Traffic Psychology and Behaviour* 76 (2021), 269–285. https://doi.org/10.1016/j.trf.2020.11.014

[95] Rebecca Wiczorek and Janna Protzak. 2022. The Impact of Visual and Cognitive Dual-Task Demands on Traffic Perception During Road Crossing of Older and Younger Pedestrians. *Frontiers in Psychology* 13 (2022), 775165. https://doi.org/10.3389/fpsyg.2022.775165

[96] Marc Wilbrink, Manja Nuttelmann, and Michael Oehl. 2021. Scaling up Automated Vehicles' EHMI Communication Designs to Interactions with Multiple Pedestrians – Putting EHMIs to the Test. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) *(AutomotiveUI '21 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 119–122. https://doi.org/10.1145/3473682.3480277

[97] George Yannis, Eleonora Papadimitriou, and Athanasios Theofilatos. 2013. Pedestrian gap acceptance for mid-block street crossing. *Transportation Planning and Technology* 36, 5 (2013), 450–462. https://doi.org/10.1080/03081060.2013.818274

[98] Xiangling Zhuang and Changxu Wu. 2012. The safety margin and perceived safety of pedestrians at unmarked roadway. *Transportation Research Part F: Traffic Psychology and Behaviour* 15, 2 (2012), 119–131. https://doi.org/10.1016/j.trf.2011.11.005