Article

Ghost on the Windshield: Employing a Virtual Human Character to Communicate Pedestrian Acknowledgement and Vehicle Intention

Alexandros Rouchitsas *[ID] and Håkan Alm

Engineering Psychology, Luleå University of Technology, 971 87 Luleå, Sweden
* Correspondence: alexandros.rouchitsas@ltu.se

Abstract: Pedestrians base their street-crossing decisions on vehicle-centric as well as driver-centric cues. In the future, however, drivers of autonomous vehicles will be preoccupied with non-driving related activities and will thus be unable to provide pedestrians with relevant communicative cues. External human–machine interfaces (eHMIs) hold promise for filling the expected communication gap by providing information about a vehicle's situational awareness and intention. In this paper, we present an eHMI concept that employs a virtual human character (VHC) to communicate pedestrian acknowledgement and vehicle intention (non-yielding; cruising; yielding). Pedestrian acknowledgement is communicated via gaze direction while vehicle intention is communicated via facial expression. The effectiveness of the proposed anthropomorphic eHMI concept was evaluated in the context of a monitor-based laboratory experiment where the participants performed a crossing intention task (self-paced, two-alternative forced choice) and their accuracy in making appropriate street-crossing decisions was measured. In each trial, they were first presented with a 3D animated sequence of a VHC (male; female) that either looked directly at them or clearly to their right while producing either an emotional (smile; angry expression; surprised expression), a conversational (nod; head shake), or a neutral (neutral expression; cheek puff) facial expression. Then, the participants were asked to imagine they were pedestrians intending to cross a one-way street at a random uncontrolled location when they saw an autonomous vehicle equipped with the eHMI approaching from the right and indicate via mouse click whether they would cross the street in front of the oncoming vehicle or not. An implementation of the proposed concept where non-yielding intention is communicated via the VHC producing either an angry expression, a surprised expression, or a head shake; cruising intention is communicated via the VHC puffing its cheeks; and yielding intention is communicated via the VHC nodding, was shown to be highly effective in ensuring the safety of a single pedestrian or even two co-located pedestrians without compromising traffic flow in either case. The implications for the development of intuitive, culture-transcending eHMIs that can support multiple pedestrians in parallel are discussed.

Keywords: external human–machine interfaces; autonomous vehicles; vehicle-to-pedestrian communication; traffic safety; gaze direction; emotional facial expressions; conversational facial expressions; neutral facial expressions

1. Introduction

1.1. Background

Interaction in traffic is officially regulated by laws and standardized communication, which is both infrastructure-based (e.g., traffic lights, traffic signs, and road surface markings) and vehicle-based (e.g., turn signals, hazard lights, and horns); still road users routinely turn to informal communication to ensure traffic safety and improve traffic flow, especially in ambiguous situations where right-of-way rules are unclear and dedicated infrastructure is missing [1,2]. Acknowledgement of other road users, intention communication, priority negotiation, and deadlock resolution are often facilitated by informal

communicative cues, such as eye contact, nodding, and waving, that are casually exchanged between drivers, motorcyclists, cyclists, and pedestrians [3,4].

In the future, however, drivers of autonomous (Level 5) vehicles will be preoccupied with non-driving related activities and will thus be unable to provide other road users with relevant communicative cues [5,6]. This development will prove particularly challenging for pedestrians, because even though they base their street-crossing decisions mainly on information provided by vehicle kinematics, such as speed and acceleration [7–9], they consider information provided by the driver too, such as the direction of their gaze and their facial expression, which has been shown to apply to different cultural contexts, such as France, China, the Czech Republic, Greece, the Netherlands, and the UK [4,10–19]. To make matters worse, previous research has shown that pedestrians tend to underestimate vehicle speed and overestimate the time at their disposal to attempt a safe crossing [20–22]. External human–machine interfaces (eHMIs), i.e., human–machine interfaces that utilize the external surface and/or the immediate surroundings of the vehicle, hold promise for filling the expected communication gap by providing pedestrians with information about the current state and future behaviour of an autonomous vehicle, as well as its situational awareness and cooperation capabilities, to primarily ensure pedestrian safety and improve traffic flow, but also to promote public acceptance of autonomous vehicle technology [23–30]. Evaluations of numerous eHMI concepts in the context of controlled studies where the participants took on the role of the pedestrian, have found interactions with vehicles that are equipped with a communication interface to be more effective and efficient and to be perceived as safer and more satisfactory compared to interactions with vehicles that are not equipped with an interface and relevant to street crossing information is provided solely by vehicle kinematics [31–48]; however see [49–51].

Although there is no consensus among researchers yet as far as optimal physical and functional eHMI characteristics are concerned, there is general agreement that the communication should be allocentric, i.e., refer exclusively to the circumstances of the vehicle instead of advising or instructing the pedestrian to act, while the provided information should be clear and relevant to the task at hand without being overwhelming or distracting [30,52]. Importantly, the ideal interface would be easily—if not intuitively—comprehensible, as comprehensibility can directly affect its effectiveness, i.e., its ability to bring about the desired result, and its efficiency, i.e., its ability to do so cost-effectively [23,53]. Previous work has shown that interface concepts employing textual message coding are easily comprehensible [34,36,41,54–57]. However, the language barrier detracts from their effectiveness and efficiency at the local level, as they can only communicate to those pedestrians who speak the language, and also limits their potential for marketability at the global level [23,30]. Interface concepts employing pictorial message coding, on the other hand, manage to transcend cultural constraints by employing widely recognized traffic symbols. Nonetheless, said symbols are visual representations of advice or instruction directed at the pedestrian, which is rather problematic due to liability issues [52]. Interestingly, interface concepts employing abstract message coding in the form of light patterns have been the most popular approach in eHMI research and development, even though they lack in comprehensibility and require explanation and training due to presenting pedestrians with the challenge of establishing new associations that are grounded in neither their experience of social interaction nor their experience of traffic interaction [31,35,41,47,48,51,58].

Interface concepts that employ anthropomorphic message coding, i.e., elements of human appearance and/or behaviour, on the other hand, tap into pedestrians' experience to communicate pedestrian acknowledgement and vehicle intention. Accordingly, Chang et al. [32] evaluated an interface concept where the headlights—serving as the "eyes" of the vehicle—turn and look at the pedestrian to communicate acknowledgement and intention to yield. Importantly, instead of being provided with information about the rationale behind the concept, the participants were required to make sense of it unaided. The results showed that the interface led to faster street-crossing decisions and higher perceived safety compared to the baseline condition (a vehicle without an interface). The findings,

however, from all other studies that have evaluated concepts employing eyes (direct gaze; averted gaze; closed eyes), facial expressions (smile; sad expression; neutral expression), and hand gestures (waving a pedestrian across the street) have been discouraging with respect to their effectiveness and efficiency [34,37,38,41,42,59,60], despite anthropomorphic autonomous vehicle-to-pedestrian communication having been lauded as a promising interaction paradigm [61]. According to Wickens et al. [62], pictorial realism and information access cost can affect the effectiveness and efficiency of a visual display. Therefore, one possible explanation for the discouraging results could be that pictorial realism was heavily compromised in these studies, given that mere abstractions of human-like elements were employed, namely, schematic faces and stand-alone hand replicas, which do not resemble in the slightest the actual communicative cues that humans experience and act on in the context of real-world social and traffic interaction. In addition to that, with the exception of Alvarez et al. [59] and Holländer et al. [42], the relevant cues were noticeably mispositioned, considering that they were presented on either the headlights, radiator grille, hood, or roof, of the vehicle, potentially leading to an increased information access cost, as pedestrians normally expect said cues to be detected at the location of the windshield.

Virtual human characters (VHCs) have long been established as tools for studying socio-perceptual and socio-cognitive processes in neurotypical and neurodiverse populations [63–69], on account of being perceived comparably to humans [70] and evoking a sense of social presence, i.e., a sense of being with another [71]. They have also been extensively employed in the field of affective computing [72,73], both as tools for studying human–computer interactions [74–79] and as end-solutions for various real-world applications in the domains of business, education, entertainment, and health [80]. For all their realism and manipulability, however, when it comes to the field of eHMIs, VHCs have so far been employed as an end-solution in only one concept. More specifically, Furuya et al. [81] evaluated an interface concept that employs a full-sized VHC sitting in the driver's seat, which either engages in eye contact with the pedestrian or keeps looking straight ahead, along the road. The participants were tasked with crossing the street at an unsignalized crosswalk while an autonomous vehicle equipped with the interface was approaching. Even though the results showed no effect of gaze direction on street-crossing efficiency or perceived safety, the vast majority of the participants preferred the virtual driver that made eye contact to both the driver with the averted gaze and the baseline condition (the vehicle without an interface), while some even touched on the usefulness of expanding its social repertoire to include gestures.

1.2. Proposed eHMI Concept

In the same spirit as Furuya et al. [81], we present an anthropomorphic eHMI concept where a VHC—male or female, displayed at the center of the windshield—engages in eye contact with the pedestrian to communicate acknowledgement. Typical driver eye-scanning behaviour in traffic consists of saccades, i.e., rapid eye movements that facilitate the detection of entities of interest in the environment, such as pedestrians; fixations, i.e., a focused gaze on an identified entity of interest; and smooth pursuit, i.e., maintaining a focused gaze in the event of a moving entity of interest [82–84]. A direct gaze is a highly salient social stimulus [85,86] that captures attention [87,88] and activates self-referential processing, i.e., the belief that one has become the focus of another's attention [89,90]. Therefore, in the context of traffic interaction, a direct gaze operates as a suitable cue for communicating pedestrian acknowledgement.

Interestingly, the VHC in the proposed concept also produces facial expressions—emotional, conversational, and neutral—to communicate vehicle intention. Previous work has shown that our distinctive ability to mentalize, i.e., to attribute mental states such as desires, beliefs, emotions, and intentions to another, is essential for accurately interpreting or anticipating their behaviour [91–93]. Emotional facial expressions are primarily regarded as the observable manifestation of the affective state of the expresser [94]. They do, however, also provide the perceiver with information about the expresser's cognitive state and

intentions [95,96]. For instance, in the context of social interaction, a Duchenne smile—the widely recognized across cultures facial expression of happiness—communicates friendliness or kindness [97], whereas an angry expression—the widely recognized across cultures facial expression of anger—communicates competitiveness or aggressiveness [98], and a surprised expression—the widely recognized across cultures facial expression of surprise—signals uncertainty or unpreparedness [99,100]. Very similar to emotional facial expressions, conversational facial expressions, such as the nod and the head shake, are widely recognized across cultures too, and provide the perceiver with information about the expresser's cognitive state and intentions [101–103]. More specifically, in the context of social interaction, a nod communicates agreement or cooperativeness, whereas a head shake communicates disagreement or unwillingness to cooperate [104–109].

The smile has been featured in several eHMI concepts as a signifier of yielding intention on the premise that, in a hypothetical priority negotiation scenario, a driver could opt for smiling at a pedestrian to communicate relinquishing the right-of-way [34,37,41,42,60,110]. Having taken context into account, the pedestrian would make the association between friendliness/kindness and said relinquishing and would thus proceed to cross the street in front of the oncoming vehicle. Following the same rationale, a driver could also opt for nodding at a pedestrian to communicate relinquishing the right-of-way [1,3,4]. In like manner, the pedestrian would make the association between agreement/cooperativeness and granting passage and would thus proceed to cross the street in front of the oncoming vehicle. Moreover, both smile and nod are social stimuli that elicit an approach tendency in the perceiver due to the presumed benefits of entering an exchange with an individual that exhibits the positive qualities assigned to said facial expressions [108,109,111,112]. In the context of traffic interaction, this tendency would translate to an intention on the part of the pedestrian to cross the street in front of the oncoming vehicle. Therefore, a smile and a nod make for suitable cues for communicating yielding intention. Accordingly, in the present study, we compared the smile to the nod with respect to their effectiveness in communicating yielding intention.

Emotional expressions have previously been featured as signifiers of non-yielding intention in only one concept. More specifically, Chang [60] evaluated a concept where yielding intention was communicated via a green happy face, whereas non-yielding intention was communicated via a red sad face. (For reasons unbeknownst, the red sad face was referred to as "a red non-smile emoticon", even though the stimulus in question clearly depicted a sad facial expression.) However, it does not make much sense for a driver—or an autonomous vehicle for that matter—to be saddened by a pedestrian, as sadness is primarily experienced due to loss while the social function of its accompanying facial expression is signaling need of help or comfort [96]. Furthermore, a sad expression is a social stimulus that elicits an approach tendency in the perceiver [113]. In real-world traffic, this tendency would translate to an intention on the part of the pedestrian to cross the street in front of the oncoming vehicle, which would be inadvisable. Therefore, a sad expression does not make for a suitable cue for communicating non-yielding intention.

In real-world traffic, it is more reasonable that a driver would be angered or surprised by a pedestrian. As a matter of fact, it has been shown that anger and surprise are among the most frequently experienced negative emotions by drivers when driving on urban roads [114]. This comes as no surprise, as urban roads are characterized by high traffic density and a plethora of situations where the reckless or inconsiderate behaviour of another road user may pose a credible danger for the driver if evasive or corrective action is not undertaken immediately [98]. Interestingly, being forced to brake because of a jaywalking pedestrian, i.e., a pedestrian that illegally crosses the street at a random uncontrolled location, has been identified as a common anger-eliciting traffic situation [115,116]. In the context of a hypothetical priority negotiation scenario, a driver could become angered by a jaywalking pedestrian and produce an angry expression to communicate their non-yielding intention. Having considered the circumstances, the pedestrian would make the association between competitiveness/aggressiveness and the refusal to grant passage and

would thus not proceed to cross the street in front of the oncoming vehicle. Similarly, a surprised expression by the driver of an oncoming vehicle could communicate a non-yielding intention to a jaywalking pedestrian. Taking context into account, the pedestrian would make the association between uncertainty/unpreparedness and the refusal to grant passage and would thus not proceed to cross the street in front of the oncoming vehicle.

Following the same rationale, a driver could also opt for shaking their head at a jaywalking pedestrian to communicate their non-yielding intention [117]. In that case, the pedestrian would make the association between disagreement/unwillingness to cooperate and denying passage and would thus not proceed to cross the street in front of the oncoming vehicle. Furthermore, all three expressions—angry, surprised, and head shake—are social stimuli that elicit an avoidance tendency in the perceiver due to the presumed costs of interacting with an individual that exhibits the negative qualities assigned to said facial expressions [99,100,108,109,111,118–121]. In the context of traffic interaction, this tendency would translate to an intention on the part of the pedestrian not to cross the street in front of the oncoming vehicle. Therefore, an angry expression, a surprised expression, and a head shake make for suitable cues for communicating non-yielding intention. Accordingly, in the present study, we compared the three facial expressions with respect to their effectiveness in communicating non-yielding intention.

The neutral expression has previously acted complementarily to the smile in eHMI research and development with respect to signifying non-yielding as well as cruising intention, i.e., the intention to continue driving in automated mode [37,41,42,110]. A neutral expression, i.e., an alert but devoid of emotion, passport-photo-like expression, is a social stimulus of neutral emotional valence, which neither readily provides information about the expresser's cognitive state or intentions nor readily elicits any action tendency in the perceiver [96,122,123]. However, in the context of a "smile-neutral" dipole, it makes sense that it be employed to signify non-yielding or cruising intention, as a pedestrian is unlikely to attempt to cross the street in front of an oncoming vehicle without receiving some form of confirmation first, sticking instead to a better-safe-than-sorry strategy rooted in ambiguity aversion [124].

Nevertheless, the neutral expression is a static facial expression that does not hint at any social—or traffic-related for that matter—interaction potential. In the proposed concept, automated mode is communicated by default via displaying the VHC on the windshield. Consequently, there is a high chance that the VHC maintaining a neutral expression will be perceived as merely an image displayed on the windshield for gimmicky purposes rather than a fully functional interface, depriving bystanders of the opportunity to form appropriate expectations and fully benefit from it. Very similar to the neutral expression, the cheek puff, i.e., making one's cheeks larger and rounder by filling them with air and then releasing it, is a facial expression of neutral emotional valence that is further considered socially irrelevant and unindicative of mental effort [63,125–127]. By extension, in the context of a "smile-cheek puff" dipole, it makes sense that it be employed to signify non-yielding or cruising intention too, as a pedestrian is more likely to stick to the above better-safe-than-sorry strategy than be convinced by a meaningless facial expression—repeatedly produced by the VHC—to cross the street in front of an oncoming vehicle. However, given that the cheek puff, unlike the neutral expression, is a dynamic facial expression, it comes with the added advantage of emphasizing the possibility of interaction with the interface. Moreover, incorporating idle-time behaviour in the social repertoire of the VHC is expected to add to its behavioural realism and positively affect its likeability and trustworthiness [77]. Accordingly, in the present study, we compared the neutral expression to the cheek puff with respect to their effectiveness in communicating cruising intention.

Dey et al. [52] have proposed a taxonomy for categorizing existing eHMI concepts across 18 parameters according to their physical and functional characteristics to enable effective evaluations and meaningful comparisons and guide the development of future implementations. For the full profile of our concept according to said work, see Appendix A.

1.3. Aim and Hypotheses

The aim of the study was to evaluate the effectiveness of the proposed concept in supporting appropriate street-crossing decision making—i.e., in preventing a pedestrian from attempting to cross the street in front of a non-yielding or cruising vehicle, without letting them pass on an opportunity to cross the street in front of a yielding vehicle— to ensure pedestrian safety without compromising traffic flow. The effectiveness of the concept was evaluated in the context of a monitor-based laboratory experiment where the participants performed a self-paced crossing intention task and their accuracy in making appropriate street-crossing decisions was measured. No explanation of the rationale behind the concept was provided beforehand to ensure intuitive responses. We hypothesized that in the "non-yielding/direct gaze" and the "cruising/direct gaze" conditions, the participants would report an intention not to cross the street. We also hypothesized that in the "yielding/direct gaze" condition the participants would report an intention to cross the street. Accuracy in the "non-yielding/direct gaze" and the "cruising/direct gaze" conditions is indicative of effectiveness in ensuring the safety of a single pedestrian intending to cross the street, whereas accuracy in the "yielding/direct gaze" condition is indicative of effectiveness in not compromising traffic flow in the case of a single pedestrian intending to cross the street [128].

Furthermore, we manipulated the participants' state of self-involvement by adding a second level to gaze direction, namely, "averted gaze" [63], to evaluate the effectiveness of the proposed concept in ensuring safety without compromising traffic flow in the case of two co-located pedestrians sharing the same intention to cross the street. (A highly entertaining rendition of a similar traffic scenario can be found here: https:// www.youtube.com/watch?v=INqWGr4dfnU&t=48s&ab_channel=Semcon, accessed on 21 April 2022). In the "averted gaze" condition, the VHC was clearly looking to the participant's right instead of directly at them [129,130], implying that someone standing next to them was the focus of its attention and the intended recipient of its communication efforts. There was, however, no explicit mention of a second, imperceivable from the participant's point of view, co-located pedestrian in the virtual scene, to ensure intuitive responses. We hypothesized that, due to the implied physical proximity, the participants would assume that the information about vehicle intention, although clearly directed at an unfamiliar roadside "neighbour", would apply to them too. Accordingly, in the "non-yielding/averted gaze" and the "cruising/averted gaze" conditions, the participants were expected to report an intention not to cross the street, whereas in the "yielding/direct gaze" condition, they were expected to report an intention to cross the street. Accuracy in the "non-yielding/averted gaze" and the "cruising/averted gaze" conditions is indicative of effectiveness in ensuring the safety of two co-located pedestrians intending to cross the street, whereas accuracy in the "yielding/averted gaze" condition is indicative of effectiveness in not compromising traffic flow in the case of two co-located pedestrians intending to cross the street [128].

2. Method

2.1. Design

A 2 × 2 × 7 within-subject design was employed, with the factors "virtual human character gender" (male; female), "gaze direction" (direct; averted), and "facial expression" (angry; surprised; head shake; neutral; cheek puff; smile; nod), serving as independent variables. Crossing the three factors yielded a total of 28 experimental conditions.

2.2. Stimuli

Twenty-eight 3D animated sequences (30 fps, 1080 × 1028 px) were developed in Poser Pro 11 (Bondware Inc.) and rendered using the Firefly Render Engine (see Appendix B). Poser figure "Rex" served as the male VHC and Poser figure "Roxie" served as the female VHC. The VHCs were shot close-up against a white background and presented in full-screen mode. Their head sizes roughly equaled that of an adult human head at a distance

of 80 cm (Rex: h = 31.5 cm, w = 20.5 cm; Roxie: h = 30.5 cm, w = 22.5 cm). In each sequence, a VHC was presented with either direct (center/0°) or averted (left/8°) gaze, bearing a neutral expression initially (200 ms) and then producing a facial expression (1134 ms: onset-peak-offset), before returning to bearing the neutral expression (100 ms). In the "neutral expression" condition, the VHC maintained a neutral expression throughout the duration of the sequence (1434 ms). All emotional expressions were of high intensity and designed to match the prototypical facial expression of the target emotion. More specifically, the Duchenne smile comprised raised cheeks, raised upper lip, upward-turned lip corners, and parted lips, i.e., the prototypical facial muscle configuration of happiness [113,131–133]. The angry expression comprised lowered, drawn-together eyebrows, raised upper eye-lid, tightened lower eyelid, and tightly pressed lips, i.e., the prototypical facial muscle configuration of anger [113,131]. The surprised expression comprised raised eyebrows, raised upper eyelid, and dropped jaw, i.e., the prototypical facial muscle configuration of surprise [113]. The neutral expression comprised relaxed overall facial musculature and lightly closed lips [123,134]. The cheek puff consisted of overly curving the cheeks outward by filling them with air, and then releasing it, executed twice while coupled with a neutral expression [63,126,127]. The nod consisted of a rigid head movement along the sagittal plane (center/0°-down/4°-center/0°), executed twice while coupled with a neutral expression [108]. The head shake consisted of a rigid head movement along the transverse plane (center/0°-left/4°-center/0°), executed twice while coupled with a neutral expression [108].

The stimulus set was validated in a pre-study where the participants performed a gaze direction and facial expression identification task (self-paced, fourteen-alternative forced choice). The pre-study was approved by the Umeå Research Ethics Committee (Ref: 2020-00642). Participant recruitment was based on convenience sampling. Fifteen participants (9 male, 6 female; mean age = 27.3 years, SD = 4.4 years) were recruited via flyers, email, social media, and personal contacts. All the participants provided written informed consent. Participation in the pre-study was voluntary and the participants could withdraw at any time without providing further explanation, however, none did so. As compensation, each participant received a lunch coupon redeemable at a local restaurant and a book of their choosing from the Engineering Psychology Group book collection. All the participants reported normal or corrected-to-normal vision. All the participants were right-handed. Seven participants were Swedish nationals, 2 Indian, 1 Bulgarian, 1 Kosovar, 1 Egyptian, 1 Mexican, 1 Iraqi, and 1 Iranian. Seven participants had an educational and/or professional background in the social sciences, 6 in engineering and technology, 1 in the medical and health sciences, and 1 other. Mean self-reported proficiency in the English language (scale of 1–10) was 7.7 (SD = 1.5; min. = 5, max. = 10). In each trial, after a fixation point appearing at the location that coincided with the center of the interpupillary line of the VHC (1000 ms), the participants were presented with one of the sequences. Then, they were presented with a list of 14 labels (VHC gender was collapsed), e.g., "Looking to my right while smiling" and "Looking at me with a neutral expression". Their task was to select the label that most accurately described the behaviour the VHC had just demonstrated. The participants could replay the sequence, if necessary, by clicking on a virtual "REPLAY" button. After responding, the participants would click on a virtual "NEXT" button to progress to the next trial. A blank white screen was presented for 1000 ms before the start of the next trial. Each sequence was presented 8 times, yielding a total of 224 experimental trials. Each sequence was presented once in randomized order before any sequence could be presented again in randomized order. Prior to the actual task, the participants completed a round of 7 practice trials, albeit with pre-selected sequences, to ensure that all levels of the three factors were represented at least once. No feedback on performance was provided during the procedure, which lasted about 45 min. Overall mean accuracy in the task was 0.981. Data from one participant were excluded from further analyses due to low mean accuracy (>2 SDs below the overall mean). The mean accuracy for the "direct gaze" and "averted gaze" conditions was 0.994 and 0.986, respectively. Mean accuracy was 0.993 for

the angry expression, 0.998 for the surprised expression, 0.987 for the head shake, 0.984 for the neutral expression, 0.996 for the cheek puff, 0.996 for the smile, and 0.975 for the nod. Mean accuracies in the "male VHC" and "female VHC" conditions were 0.992 and 0.987, respectively. After the identification task, the participants rated the VHCs on 7 dimensions (5-point Likert scale: "not at all" = 0, "slightly" = 1, "moderately" = 2, "very" = 3, and "extremely" = 4). Mean scores were 1.6 for realism, 1.4 for familiarity, 1.9 for friendliness, 1.2 for attractiveness, 1.9 for trustworthiness, 0.9 for threateningness, and 2.0 for naturalness.

2.3. Apparatus

The task was run on a Lenovo ThinkPad P50s (Intel® Core™ i7-6500U CPU @ 2.5 GHz; 8 GB RAM; Intel HD Graphics 520; Windows 10 Education; Lenovo, Beijing, China). The stimuli were presented on a 24" Fujitsu B24W-7 LED monitor. The participants responded via a Lenovo MSU1175 wired USB optical mouse. Stimuli presentation and data collection were controlled by E-Prime 3.0 (Psychology Software Tools Inc., Sharpsburg, PA, USA).

2.4. Procedure

The experiment was conducted in a separate, quiet, and well-lit room at the researchers' institution. Upon arrival, the participants were seated and asked to read and sign a consent form including information about study aim, procedure, data handling, participation risks, and compensation. The participants then provided the experimenter with demographic information. After assuming a standing position in front of the screen—at approximately 80 cm away from the screen, whose angle was adjusted so that the VHC met each individual participant's eyeline (Figure 1)—they were presented with an informational passage alongside a visualization of the proposed eHMI concept (Figure 2), for familiarization purposes. The passage read: "Interaction between drivers and pedestrians is often assisted by gaze direction, facial expressions, and gestures. In the near future, however, users of autonomous cars will not have to pay attention to the road; they will read, work, watch movies, or even sleep, while the car takes them where they want to go. Therefore, to make traffic interaction safer and easier, advanced communication systems will have to be developed that inform pedestrians about whether an autonomous car has detected them and what it plans on doing next. Communication systems that utilize the windshield area of an autonomous car have been proposed as a possible solution to the problem pedestrians will soon face. For this reason, we have developed an LED-display system where a 2D avatar—substituting for the human driver—is displayed on the windshield, to communicate pedestrian detection and car intention. You are here to test our system." (As English was expected to be a second language for most participants, we chose to present the introductory information in as plain English as possible. As a result, the words "vehicle", "acknowledgement", and "interface", were replaced by "car", "detection", and "system", respectively. Moreover, the 3D VHCs were referred to as "2D" VHCs to manage expectations regarding depth of field, given they were presented in a decontextualized fashion.)

While maintaining the standing position, which was employed to facilitate the adoption of pedestrians' perspective, the participants performed a crossing intention task (self-paced, two-alternative forced choice). (Previous research has shown that performance in a primary cognitive task is not affected by assuming a standing position [135,136].) In each trial, after a fixation point appearing at the location that coincided with the center of the interpupillary line of the VHC (1000 ms), the participants were presented with one of the sequences. Then, they were presented with two choices, namely, "Cross" and "Not cross", and the following instructions: "Imagine you are a pedestrian about to cross a one-way street at a random uncontrolled location when you see an autonomous car approaching from the right, that is equipped with our communication system. What would you do if the avatar on the windshield demonstrated the same behaviour as the virtual character: Would you cross the street or not cross the street?" (Following the recommendation of Kaß et al. [137], we had the pedestrian intending to cross the street in front of a single oncoming vehicle at a random uncontrolled location to ensure that the participants' responses were

unaffected by other traffic and expectations regarding right-of-way.) The participants could replay the sequence, if necessary, by clicking on a virtual "REPLAY" button. After responding, the participants would click on a virtual "NEXT" button to progress to the next trial. A blank white screen was presented for 1000 ms before the start of the next trial. Each sequence was presented 10 times, yielding a total of 280 experimental trials. Each sequence was presented once in randomized order before any sequence could be presented again in randomized order. Prior to the actual task, the participants completed a round of 7 practice trials, albeit with pre-selected sequences, to ensure that all levels of the three factors were represented at least once. No feedback on performance was provided during the procedure, which lasted about 45 min.

Figure 1. Experiment set-up. (Photo taken for demonstration purposes. The participants responded via mouse in the actual task.).
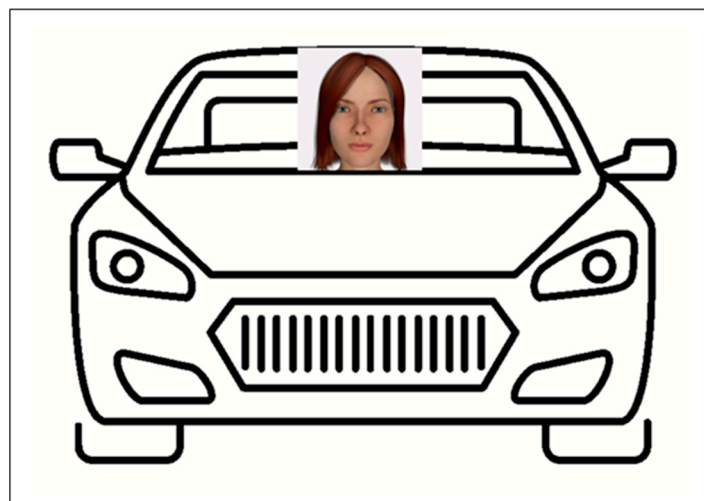
Figure 2. Visualization of the proposed eHMI concept.

2.5. Dependent Variable

Accuracy (total correct trials/total trials) for crossing intention responses was measured. "Not cross" was coded as the correct response for the angry expression, the surprised expression, the head shake, the neutral expression, and the cheek puff, regardless of gaze direction and VHC gender condition. "Cross" was coded as the correct response for the smile and the nod, regardless of gaze direction and VHC gender condition. Following the recommendation of Kaß et al. [137], the absolute criterion for effectiveness was set at 0.85 correct in each experimental condition.

2.6. Participants

The study was approved by the Umeå Research Ethics Committee (Ref: 2020-00642). Participant recruitment was based on convenience sampling. Thirty participants (18 male, 12 female; mean age = 33.1 years, SD = 11.9 years) were recruited via flyers, email, social media, and personal contacts. None had participated in the stimuli validation pre-study. All the participants provided written informed consent. Participation in the study was voluntary and the participants could withdraw at any time without providing further explanation, although none did so. As compensation, each participant received a lunch coupon redeemable at a local restaurant and a book of their choosing from the Engineering Psychology Group book collection.

All the participants reported normal or corrected-to-normal vision. Twenty-seven participants were right-handed. Fourteen participants were Swedish nationals, 4 Indian, 2 German, 1 Greek, 1 French, 1 Egyptian, 1 Yemeni, 1 Venezuelan, 1 Brazilian, 1 Colombian, 1 Pakistani, 1 Japanese, and 1 Chinese. Twenty-three participants had an educational and/or professional background in engineering and technology, 5 in the social sciences, 1 in the medical and health sciences, and 1 in the arts. Twenty-five participants assumed the role of a pedestrian on a daily and 5 on a weekly basis. Twenty-eight participants had a driving license. Mean period of residence in Sweden was 17.6 years (SD = 17.5; min. = 1, max. = 67). Mean self-reported proficiency in the English language (scale of 1–10) was 8.2 (SD = 1.3; min. = 6, max. = 10).

2.7. Data Analysis

We tested the effects of VHC gender, gaze direction, and facial expression on the dependent variable for each vehicle intention separately. For non-yielding vehicles, a 2 (male; female) $\times$ 2 (direct; averted) $\times$ 3 (angry expression; surprised expression; head shake) repeated measures analysis of variance (ANOVA) was conducted. For cruising vehicles, a 2 (male; female) $\times$ 2 (direct; averted) $\times$ 2 (neutral expression; cheek puff) repeated measures ANOVA was conducted. For yielding vehicles, a 2 (male; female) $\times$ 2 (direct; averted) $\times$ 2 (smile; nod) repeated measures ANOVA was conducted. Data were processed in Excel (Microsoft) and analyzed in SPSS Statistics 28 (IBM). We used Mauchly's test to check for violations of the assumption of sphericity, and degrees of freedom were adjusted using the Greenhouse–Geiser estimate of sphericity where necessary [138]. We also used Bonferroni-corrected pairwise comparisons where necessary.

3. Results

3.1. Non-Yielding Vehicles

The mean accuracy in the direct gaze condition was 0.962 for the angry expression, 0.997 for the surprised expression, and 0.997 for the head shake (Table 1). The mean accuracy in the averted gaze condition was 0.965 for the angry expression, 0.992 for the surprised expression, and 0.998 for the head shake. The mean accuracy in the male VHC and female VHC conditions was 0.984 and 0.986, respectively. According to the results, all three facial expressions communicated vehicle intention effectively to a single pedestrian intending to cross the street in the presence of a non-yielding vehicle. Additionally, the proposed concept communicated vehicle intention effectively to two co-located pedestrians

intending to cross the street in the presence of a non-yielding vehicle. The VHC gender did not differentially affect accuracy.

These impressions were confirmed by a 2 × 2 × 3 repeated measures ANOVA that revealed a non-significant main effect of VHC gender, $F(1, 29) = 0.659$, $p = 0.423$, $\eta_p^2 = 0.022$; a non-significant main effect of gaze direction, $F(1, 29) = 0.000$, $p > 0.999$, $\eta_p^2 = 0.000$; and a non-significant main effect of expression, $F(1.006, 29.173) = 1.001$, $p = 0.326$, $\eta_p^2 = 0.033$. All the interactions were non-significant.

Table 1. Mean accuracy (standard error) for crossing intention according to design rationale. Absolute criterion for effectiveness was set at 0.85 correct in each experimental condition.

|		Facial Expression						
		Angry	Surprised	Head Shake	Neutral	Cheek Puff	Smile	Nod
Gaze Direction	Direct	0.962 (0.033)	0.997 (0.002)	0.997 (0.002)	0.823 (0.066)	0.868 (0.06)	0.76 (0.075)	0.963 (0.027)
	Averted	0.965 (0.033)	0.992 (0.004)	0.998 (0.002)	0.822 (0.064)	0.867 (0.061)	0.73 (0.076)	0.937 (0.034)
		Non-yielding			Cruising		Yielding	
		Vehicle Intention						

3.2. Cruising Vehicles

The mean accuracy in the direct gaze condition was 0.823 for the neutral expression and 0.868 for the cheek puff (Table 1). The mean accuracy in the averted gaze condition was 0.822 for the neutral expression and 0.867 for the cheek puff. The mean accuracies in the male VHC and female VHC conditions were 0.843 and 0.847, respectively. According to the results, both facial expressions communicated vehicle intention to a single pedestrian intending to cross the street in the presence of a cruising vehicle comparably well, however, only the cheek puff did so effectively. Additionally, the proposed concept communicated vehicle intention effectively to two co-located pedestrians intending to cross the street in the presence of a cruising vehicle only in the cheek puff condition. The VHC gender did not differentially affect accuracy.

These impressions were confirmed by a 2 × 2 × 2 repeated measures ANOVA that revealed a non-significant main effect of VHC gender, $F(1, 29) = 0.563$, $p = 0.459$, $\eta_p^2 = 0.019$; a non-significant main effect of gaze direction, $F(1, 29) = 0.028$, $p = 0.869$, $\eta_p^2 = 0.001$; and a non-significant main effect of expression, $F(1, 29) = 0.291$, $p = 0.594$, $\eta_p^2 = 0.010$. All the interactions were non-significant.

3.3. Yielding Vehicles

The mean accuracy for the direct gaze condition was 0.76 for the smile and 0.963 for the nod (Table 1). The mean accuracy for the averted gaze condition was 0.73 for the smile and 0.937 for the nod. The mean accuracies for the male VHC and female VHC conditions were 0.847 and 0.847, respectively. According to the results, the nod greatly outperformed the smile in communicating vehicle intention effectively to a single pedestrian intending to cross the street in the presence of a yielding vehicle. Additionally, the proposed concept communicated vehicle intention effectively to two co-located pedestrians only in the nod condition. The VHC gender did not differentially affect accuracy.

These impressions were confirmed by a 2 × 2 × 2 repeated measures ANOVA that revealed a non-significant main effect of VHC gender, $F(1, 29) = 0.000$, $p > 0.999$, $\eta_p^2 = 0.000$, and a non-significant main effect of gaze direction, $F(1, 29) = 1.094$, $p = 0.304$, $\eta_p^2 = 0.036$. However, the 2 × 2 × 2 repeated measures ANOVA revealed a significant main effect of expression: $F(1, 29) = 8.227$, $p = 0.008$, $\eta_p^2 = 0.221$ (Figure 3). All the interactions were non-significant.
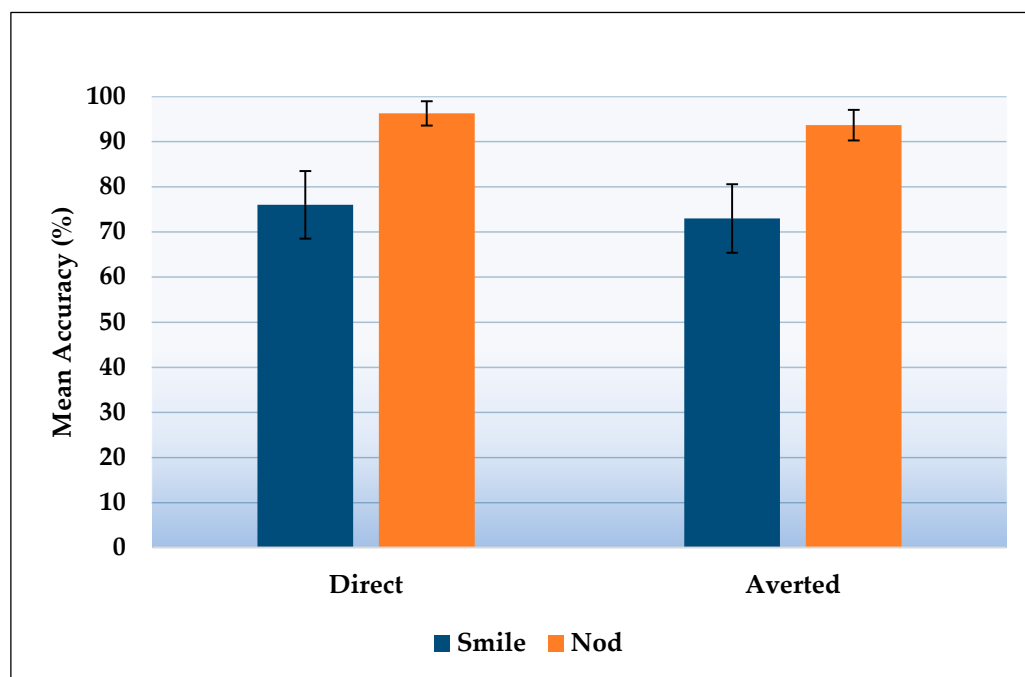
Figure 3. Mean accuracy (%) for crossing intention in the presence of yielding vehicles for direct and averted gaze. Error bars represent standard error.

4. Discussion

4.1. Findings

The effectiveness of the angry expression, the surprised expression, and the head shake in communicating non-yielding intention was to be expected, given that these widely recognized facial expressions communicate competitiveness/aggressiveness, uncertainty/unpreparedness, and disagreement/uncooperativeness, respectively, while eliciting an avoidance tendency in the perceiver. Moreover, as hypothesized, in the averted gaze condition, due to the implied physical proximity, the information about vehicle intention—although clearly intended for someone else—did support the participants decide appropriately when interacting with a non-yielding vehicle. Therefore, an implementation of our concept where non-yielding intention is communicated via either an angry expression, a surprised expression, or a head shake would be highly effective in ensuring the safety of a single pedestrian (M = 98.5%) or even two co-located pedestrians (M = 98.5%) intending to cross the street in the presence of a non-yielding vehicle.

With respect to communicating cruising intention, the effectiveness of both the neutral expression and cheek puff was also to be expected, considering that ambiguous or meaningless facial expressions are unlikely to convince pedestrians to risk being run over in their attempt to cross the street. Much to our surprise, however, since it is the field's standard anthropomorphic signifier of denying passage, the neutral expression missed the 85% mark for effectiveness, though only by a little. Regarding the averted gaze condition, facial expression was again shown to take precedence over gaze direction and inform the participants' decisions when interacting with a cruising vehicle too. Hence, our concept proved to be effective in ensuring the safety of a single pedestrian or even two co-located pedestrians intending to cross the street in the presence of a cruising vehicle only when operating in cheek-puff mode (86.8%/86.7%). Taken together, these results suggest that an implementation of our concept where a non-yielding intention is communicated via either an angry expression, a surprised expression, or a head shake, and where cruising intention is communicated via cheek puff, would be highly effective in the context of safety-critical scenarios, i.e., in interactions with non-yielding or cruising vehicles, in both the case of a single pedestrian (M = 92.7%) and that of two co-located pedestrians (M = 92.6%).

Interestingly, the smile and the nod did not communicate yielding intention comparably well, with the nod greatly outperforming the field's go-to anthropomorphic signifier of relinquishing the right-of-way, which missed the 85% effectiveness mark by a lot. Considering that previous work has shown that both smile and nod are widely recognized facial expressions with clear associations, and that their identification rates in our pre-study were 99.6% and 97.5%, respectively, it is unlikely that this performance discrepancy is due to participant confusion—at the perceptual or cognitive level—specific to the smile condition. A more plausible explanation is that, in the context of a hypothetical negotiation scenario, a smile is more open to interpretation than a nod, as one could opt for smiling to the other party to communicate friendliness and/or kindness, all the while not intending to succumb to their demands, whereas opting for nodding would most likely indicate reaching consensus and the successful resolution of the negotiation process.

In the context of our experiment, the nod's advantage in unambiguousness over the smile manifested in a considerably higher rate of correct responses in the nod condition compared to the smile condition. Apparently, an ambiguous stimulus, such as the neutral expression, is confirmation enough for a pedestrian to make a conservative decision, namely, to not cross the street in the presence of what could still in theory be a yielding vehicle, and only risk wasting their time; however, an ambiguous stimulus, such as the smile, is not confirmation enough for a pedestrian to make a liberal decision, namely, to cross the street in the presence of what could still in theory be a non-yielding vehicle, and jeopardize their well-being. On the other hand, the nod appears to be categorical enough to lead to a decision to cross the street in the presence of what could still in theory be a non-yielding vehicle. Moreover, as hypothesized, information about vehicle intention, even when clearly directed at an unfamiliar roadside "neighbour", was also shown to support the participants decide appropriately when interacting with a yielding vehicle. Therefore, as far as effectiveness in not compromising traffic flow is concerned, an implementation of our concept where yielding intention is communicated via nod would be highly effective in both the case of a single pedestrian (96.3%) and that of two co-located pedestrians (93.7%) intending to cross the street in the presence of a yielding vehicle.

With respect to our manipulation of the participants' state of self-involvement, it is important to note that performance in the employed task was contingent neither on attention to gaze nor on gaze direction discrimination, as vehicle intention can be inferred purely from the facial expression. So, in theory, one could attend only to the lower half of the screen and still perform the task successfully, provided they would correctly identify the angry expression, the surprised expression, and the smile on the basis of information from the mouth region only [101,139]. To prevent the participants from adopting this response strategy while ensuring intuitive responses, we decided to abstain from instructing them to attend to the eye region, and instead adjust the screen angle at the start of the procedure, so that the VHC met each individual participant's eyeline, and have the fixation point appearing at the center of the interpupillary line of the VHC, to implicitly guide attention at the start of each trial to the exact location of the screen the eye region would soon occupy [140,141]. Even though we see no advantage in adopting said response strategy—as it will not decrease the task completion time and is more demanding attention-wise, given that the face is holistically processed at a glance [142], whereas the gaze requires focused attention to be ignored [143]—we cannot rule out the possibility that it may have been adopted by someone. In any case, we readily accept this risk, as we believe that the benefit of ensuring intuitive responses greatly outweighs the harm induced from a few compromised data sets.

4.2. Implications

Although eHMI concepts employing abstract or textual message coding do support pedestrians in their street-crossing decision-making, they typically require either explanation and training or cultural adaptation. Furthermore, their implementation presupposes the introduction of additional stimuli in an already overloaded traffic environment, which

will only serve to further confuse and frustrate road users, and possibly negatively affect the acceptance of autonomous vehicle technology [30]. The proposed concept, on the other hand, manages to be effective while bypassing these concerns due to employing anthropomorphic message coding instead, and tapping into pedestrians' experience of social and traffic interaction. More specifically, employing salient and widely recognized communicative cues with clear associations resulted in an interface concept that is both intuitively comprehensible and culture-transcending. This is consistent with Singer et al. [46], who have provided evidence for the effectiveness of a uniform eHMI concept in communicating vehicle intention across different cultural contexts, namely, in China, South Korea, and the USA. Considering that the perceived ease of use is a main determinant of automation acceptance [144], the simplicity of our concept is also likely to positively affect the acceptance of autonomous vehicle technology. Lastly, the VHC displayed on the windshield will not place any additional mental burden on road users, as it will occupy the physical space otherwise occupied by the driver of the vehicle [145].

Nonetheless, some facial expressions fared considerably better in embodying our design rationale than others. More specifically, in the case of cruising intention, the neutral expression's failure to reach the 85% effectiveness mark suggests that its employment as a signifier of denying passage will have to be revisited. Minimizing ambiguousness is essential when designing for intuitive communication in the context of safety-critical scenarios [146,147], so opting instead for a meaningless facial expression, which was shown to breed more reluctance in pedestrians and will draw more attention to it on account of being dynamic [93], is an alternative worth exploring. Similarly, the smile's failure to reach the 85% effectiveness mark in communicating yielding intention suggests that its employment as a signifier of granting passage will also have to be revisited. Even though it is not a matter of traffic safety in this case but of traffic flow, traffic flow is still the other main ambition of the traffic system and minimizing ambiguousness by opting for the nod instead has the potential to positively affect its overall performance.

Regarding the employment of VHCs, it has been argued that referencing current forms of communication in traffic will be key for any novel eHMI concept if it is to be introduced into the traffic system without disrupting regular traffic operation [23]. From a pedestrian's point of view, a VHC displayed on the windshield, which employs its gaze direction and facial expressions to communicate pedestrian acknowledgement and vehicle intention, respectively, is a direct reference to their common experience of turning to informal communication to facilitate interaction with the driver of an oncoming vehicle. (In our case, as the vehicle is approaching from the right side, the right-side window is also an area of interest for the pedestrian. However, given that the 3D VHCs were presented in a decontextualized fashion, it would have made no difference to the participant to distinguish between the windshield and right-side window during the experimental procedure. Therefore, the participants were presented with simplified instructions that made mention of only the windshield area of the vehicle.) However, an interface should also be distinctive enough to communicate driving mode effectively and dynamic enough to betray its interaction potential. In the proposed concept, automated mode is communicated by default via displaying the VHC on the windshield, while the interaction potential is communicated via the VHC puffing its cheeks repeatedly when the vehicle is cruising. All things considered, our concept seems to strike a fine balance between familiarity and novelty the field needs more of if eHMIs are to be integrated into the traffic system successfully.

Placing the effectiveness of our concept aside, there appears to be some confusion regarding the clarity with which anthropomorphic eHMI concepts employing gaze direction to communicate acknowledgement and/or vehicle intention manage to pinpoint a single pedestrian. In Dey et al. [52], the "Eyes on a Car" and "Prototype 3/visual component" evaluated in Chang et al. [32] and Mahadevan et al. [37], respectively, were classified as "Unclear Unicast", suggesting that the interfaces can address only one pedestrian at a time (unicast), albeit without clearly specifying which pedestrian (unclear). However, even though they do address only one pedestrian at a time, eye contact with that one pedestrian

is established in both concepts to communicate that they specifically have become the focus of attention and will be the intended recipient of any upcoming communication efforts on the VHC's part. While Chang et al. and Mahadevan et al. did not measure gaze direction accuracy, the mean accuracies in our pre-study in the "direct gaze" and "averted gaze" conditions were 99.4% and 98.6%, respectively. Given that humans have evolved to be very adept at gaze direction discrimination due to the adaptive value of the trait [86], we strongly believe that anthropomorphic eHMI concepts employing gaze direction to communicate pedestrian acknowledgement and/or vehicle intention should be classified as "Clear Unicast", indicating that they can clearly specify who that one pedestrian they can address at a time is. Accordingly, if an anthropomorphic eHMI concept employing gaze direction were to fail in pinpointing a single pedestrian, said failure should be attributed to the implementation of the concept being defective, and not to the theory behind it being unsubstantiated.

However, for an eHMI to be considered effective in real-world traffic, not to mention a profitable endeavour for manufacturers, supporting a single pedestrian make appropriate street-crossing decisions when interacting with a single autonomous vehicle is not sufficient. Considering the average traffic scenario is far more complex than that, as multiple road users are involved, an effective interface must also show promise of supporting more than one pedestrian at a time [148]. Previous work on pedestrian following behaviour has shown that pedestrians incorporate social information provided by their unfamiliar roadside "neighbours", namely, their crossing behaviour, into their own street-crossing decision-making formula, and are, therefore, likelier to cross the street behind another crossing pedestrian [149]. This prevalence of following behaviour in traffic has led to the dismissal of eHMI evaluation scenarios involving two co-located strangers sharing the same intention to cross the street, on the assumption that an interface could support both at the same time, without there being any need of distinguishing between them [150]. Nonetheless, given that research on whether pedestrian following behaviour in the presence of manually driven vehicles carries over to interactions with autonomous vehicles is still lacking, such a dismissal seems somewhat premature.

Recent work has opted for evaluating eHMI concepts in the context of more challenging scenarios instead, involving a second, non-collocated pedestrian sharing the same intention to cross the street. However, evidence of the potential of on-vehicle interfaces for supporting more than one pedestrian at a time is yet to be produced. More specifically, Wilbrink et al. [151] evaluated an on-vehicle concept that employs abstract message coding in the form of light patterns on the windshield to communicate vehicle intention. Even though the participants were presented with the rationale behind the concept beforehand, and the second pedestrian was constantly visible from their point of view, the mere presence of another pedestrian in the virtual scene was enough to confuse them, as evidenced by the lower reported willingness to cross the street compared to interacting with the oncoming autonomous vehicle on a one-to-one basis. In the same vein, Dey et al. [152] evaluated three eHMI concepts designed to complement a core concept by providing additional contextual information, namely, to whom, when, and where an oncoming autonomous vehicle would yield. Even though the rationale behind the concepts had been explained in full detail beforehand and the second pedestrian was visible throughout the task, the on-vehicle concepts negatively affected the interpretation of vehicle intention, as evidenced by the higher reported willingness to cross the street when the vehicle was in fact yielding to the other pedestrian compared to the baseline condition (a vehicle without an interface). Only in the case of a projection-based concept were the participants more accurate in interpreting vehicle intention, as evidenced by the lower reported willingness to cross the street when the vehicle was yielding to the other pedestrian. Nevertheless, the effectiveness of projection-based concepts in real-world traffic has been questioned due to their susceptibility to weather conditions, road surface conditions, and time of day, and exploring further the potential of on-vehicle interfaces for supporting more than one pedestrian at a time has been recommended instead [52].

To the best of our knowledge, our study is the first to provide evidence of the potential of an on-vehicle interface for supporting more than one pedestrian at a time make appropriate street-crossing decisions when interacting with an oncoming autonomous vehicle. (Joisten et al. [153] evaluated two on-vehicle eHMI concepts in single-pedestrian and three-pedestrian (co-located) scenarios. However, due to a major confound, namely, the inclusion of a crosswalk in the virtual scene in every experimental condition, their results have been rendered uninterpretable with respect to the effectiveness of the concepts due to clear expectations regarding the right-of-way.) More specifically, in the averted gaze condition, the participants experienced the interface as if they belonged to a group of two co-located pedestrians sharing the same intention to cross the street while only the other pedestrian was acknowledged by the VHC. Importantly, there had been no explanation of the rationale behind the proposed concept beforehand and no mention of a second pedestrian being present in the virtual scene. Even so, the results showed that the participants neither seriously compromised their safety by deciding too liberally, namely, by reporting an intention to cross regardless of vehicle intention, on the basis that the communication was not directed at them, nor greatly compromised traffic flow by deciding too conservatively, namely, by reporting an intention not to cross regardless of vehicle intention, on the basis that acknowledgement is necessary to proceed. Rather, they accurately inferred vehicle intention regardless of pedestrian acknowledgement, which serves as evidence of the effectiveness of an anthropomorphic eHMI concept employing gaze direction to communicate acknowledgement in supporting at least two co-located pedestrians decide appropriately when interacting with non-yielding, cruising, and yielding autonomous vehicles.

5. Considerations

5.1. Trust and Acceptance

Trust is a complex psychological state arising from one's perceived integrity, ability, and benevolence of another [154]. Similarly, trust in automation is built on the basis of information about the performance, processes, and purpose of the technology in question, and is a main determinant of automation acceptance [155]. Importantly, both too little trust in automation, i.e., distrust, and too much, i.e., overtrust, can negatively affect interactions with technology [156]. In the case of distrust, the likelihood and consequences of a malfunction are overestimated, and the technology is perceived as less capable than it is. Accordingly, a distrusting pedestrian will reluctantly delegate, if at all, responsibility for a successful interaction with an oncoming autonomous vehicle to the interface, which will negatively affect traffic flow if the vehicle does yield after all. Evidently, calibrating trust in the interface, i.e., adjusting trust to match its actual trustworthiness, is essential if it is to ever reach its full performance potential in real-world traffic [157,158]. Considering that trust mediates automation acceptance, it becomes all the more important to allay skepticism and hesitation regarding eHMIs, especially during the familiarization phase, so that the public fully benefits from autonomous vehicle technology [159–164]. Previous work has shown that anthropomorphizing an interface ensures that greater trust is placed in it [155,165–167]. Furthermore, anthropomorphizing an autonomous vehicle increases trust in its capabilities [168].

On the other hand, in the case of overtrust in automation, the likelihood and consequences of malfunction are underestimated, and the technology is perceived as more capable than it is [156]. Accordingly, an overtrusting pedestrian will carelessly delegate responsibility for a successful interaction with an oncoming autonomous vehicle to the interface, which may put them in danger if the vehicle does not yield after all. Previous research has shown that pedestrians trust an autonomous vehicle that is equipped with an eHMI more than they do a vehicle that is not [169], especially when having been provided with information about its performance, processes, and purpose prior to the initial interaction [170]. Interestingly, unlike trust in another [171], trust in eHMIs appears to be very robust, as an incident of interface malfunction will only momentarily reduce trust in the interface and the acceptance of autonomous vehicles [158,172] An anthropomorphic

eHMI is likely to increase the potential for overtrust, as previous work has shown that anthropomorphism positively affects perceived trustworthiness [173]. Moreover, a VHC displayed on the windshield of an oncoming autonomous vehicle is likely to elicit curiosity or even astonishment in bystanders, potentially distracting them in their attempt to navigate regular traffic or interact with the vehicle effectively [145]. Therefore, educating vulnerable road users to appreciate the capabilities and shortcomings of novel interaction paradigms will be key to ensuring their safety in traffic and the successful deployment and integration of autonomous vehicle technology in society.

5.2. Limitations

Certain limitations of the present study must be acknowledged, as they may have compromised the external validity of our findings [174]. Firstly, the proposed concept was evaluated in the context of a monitor-based laboratory experiment, which guarantees participants' safety and provides researchers with greater experimental control, flexibility in factor manipulation, and efficiency in prototyping, compared to a field experiment employing a physical prototype. Nevertheless, the chosen approach suffers from limited ecological validity due to the high degree of artificiality and low degree of immersion of the experiment set-up, as well as the possibility of response bias due to the lack of safety concerns. Evidently, the logical next step would be for our concept to be evaluated in the context of a virtual reality (VR)-based laboratory experiment that effectively combines the advantages of a typical laboratory experiment with the high degree of immersion of a VR environment to obtain more ecologically valid data [175,176].

Secondly, the participants in our study were presented with an overly simplistic traffic scenario to ensure that their responses were unaffected by other traffic and expectations regarding right-of-way [137]. However, this decision may have rendered our findings nongeneralizable to more complex scenarios that involve multiple road users and dedicated infrastructure. Thus, our concept should also be evaluated in the context of more elaborate scenarios to gain a fuller understanding of its potential effectiveness in real-world traffic.

Thirdly, the VHCs were presented in a decontextualized fashion to ensure that the participants' responses were unaffected by visual elements of lesser importance and information provided by vehicle kinematics [56]. Yet, this mode of stimuli presentation may have led to an overestimation of the effectiveness of the proposed concept, considering that both the salience and discriminability of the social information provided by the interface would have most likely been lower had the interface been encountered in real-world traffic due to the distance from and speed of the oncoming vehicle or due to unfavourable environmental conditions, such as poor lighting or rain. Interestingly, however, the opposite may also be true, namely, that the decontextualization may have led to an underestimation, as the accuracy would have most likely been higher had the information provided by visual elements of lesser importance and/or vehicle kinematics also been considered in the disambiguation process [177,178]. To measure how the effectiveness of our concept is affected by visual elements of lesser importance, the information provided by vehicle kinematics, or unfavourable environmental conditions, the employed VHCs should also be experienced in the context of more perceptually and cognitively demanding virtual scenes.

Fourthly, due to technical and spatial constraints, we employed a task where no locomotion was involved, as the participants indicated their street-crossing decisions simply via mouse click. Although input devices, such as response boxes, keyboards, mouses, and sliders, are common in laboratory studies, and have also been extensively utilized in eHMI research, it is evident that the odd substitution of actual forward steps with finger movements serves only to detract from the ecological validity of the findings. Hence, our concept should also be evaluated in a pedestrian simulator to gain a fuller understanding of its potential for facilitating actual crossing behaviour [179,180].

Lastly, due to time constraints, the procedure did not include an explicitation interview [181,182], potentially depriving us of the opportunity to gain valuable insight into the mental processes that guided the participants' responses. In any case, we plan on

incorporating said technique in subsequent evaluations so that participants can freely elaborate on their subjective experience when interacting with our interface.

5.3. Future Work

The aim of the present study was to evaluate the effectiveness of the proposed concept in supporting pedestrians make appropriate street-crossing decisions when interacting with non-yielding, cruising, and yielding autonomous vehicles. However, in real-world traffic, deciding appropriately cannot be the only concern; appropriate decisions should be made within a reasonable time-frame, to achieve a viable compromise between the two main ambitions of traffic safety and traffic flow [183]. Accordingly, in future work, we plan to also evaluate the efficiency of our concept, i.e., its effectiveness in supporting pedestrians make appropriate street-crossing decisions in a timely fashion.

Previous work has shown that children tend to rely entirely on the interface for their street-crossing decisions when interacting with an autonomous vehicle, altogether ignoring the information provided by vehicle kinematics [184]. For children, who are more vulnerable than the average pedestrian due to their inexperience, playfulness, and carelessness, dealing with the novelty of an anthropomorphic eHMI may prove particularly challenging [185,186]. Therefore, in future work, we plan to evaluate the effectiveness and efficiency of our concept with a group of children to assess the extent to which they could be at a disadvantage during interactions with the interface.

Furthermore, successful interaction with the proposed concept presupposes intact socio-perceptual and socio-cognitive abilities, such as gaze direction discrimination and mentalizing. Relevant research has shown that impaired communication and social interaction skills are among the defining characteristics of autism spectrum disorder [66,187,188]. Accordingly, in future work, we plan to evaluate the effectiveness and efficiency of our concept with a group of autistic individuals to assess the extent to which they could be at a disadvantage when interacting with our interface, especially considering that pedestrians with neurodevelopmental disorders are already at a higher risk of injury in traffic compared to neurotypical pedestrians [189].

6. Conclusions

An implementation of the proposed anthropomorphic eHMI concept where non-yielding is communicated via either an angry expression, a surprised expression, or a head shake; cruising is communicated via cheek puff; and yielding is communicated via nods, was shown to be highly effective in ensuring the safety of at least two co-located pedestrians in parallel without compromising traffic flow. Importantly, this level of effectiveness was reached in the absence of any explanation of the rationale behind the eHMI concept or training to interact with it successfully, as it communicates pedestrian acknowledgement and vehicle intention via communicative cues that are widely recognized across cultures. Therefore, it has been concluded that the proposed anthropomorphic eHMI concept holds promise for filling the expected communication gap between at least two co-located pedestrians sharing the same intention and an oncoming autonomous vehicle, across different cultural contexts.

Author Contributions: A.R. contributed to conceptualization, experimental design, stimuli creation, experiment building, data collection, data processing, data analysis, manuscript preparation, and manuscript revision. H.A. contributed to manuscript revision and supervision. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Profile of the Proposed eHMI Concept According to Dey et al. [52]

1. Target road user: Pedestrians.
2. Vehicle type: Passenger cars.
3. Modality of communication: Visual (anthropomorphic).
4. Colors for visual eHMIs: N/A.
5. Covered states: Non-yielding; cruising; yielding.
6. Messages of communication in right-of-way negotiation: Situational awareness; intention.
7. HMI placement: On the vehicle.
8. Number of displays: 1.
9. Number of messages: 3.
10. Communication strategy: Clear unicast.
11. Communication resolution: High.
12. Multiple road user addressing capability: Single.
13. Communication dependence on distance/time gap: No.
14. Complexity to implement: C4 (uses technology that is not yet developed or not widely available on the market).
15. Dependence on new vehicle design: No.
16. Ability to communicate vehicle occupant state/shared control: No.
17. Support for people with special needs: No.
18. Evaluation of concept: Yes

- Time of day: Unspecified.
- Number of simultaneous users per trial: 1.
- Number of simultaneous vehicles per trial: 1.
- Method of evaluation: Monitor-based laboratory experiment.
- Weather conditions: Unspecified.
- Road condition: Unspecified.
- Sample size: 30.
- Sample age: M = 33.1 years, SD = 11.9 years.
- Measures: Performance (accuracy).

Appendix B. Stills Taken from the 3D Animated Sequences
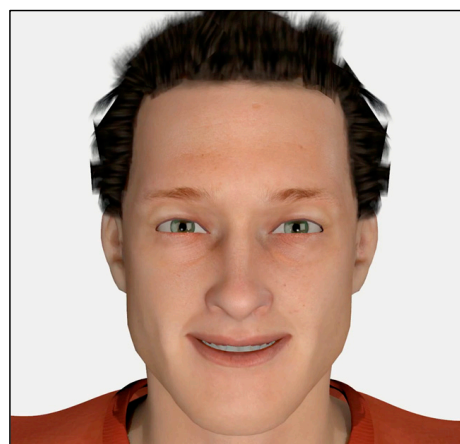
Figure A1. Rex looking directly at the participant while smiling.
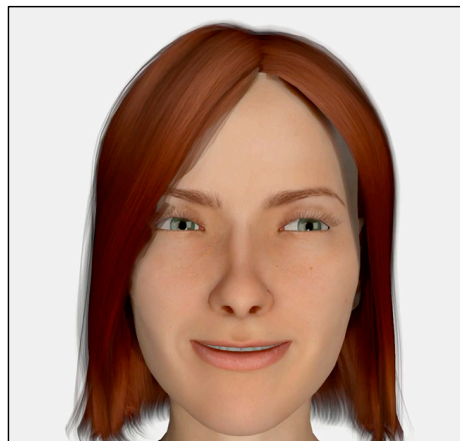
Figure A2. Rex looking to the participant's right while smiling.

Figure A3. Roxie looking directly at the participant while smiling.
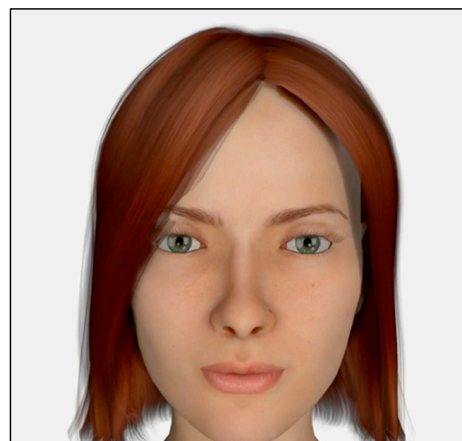

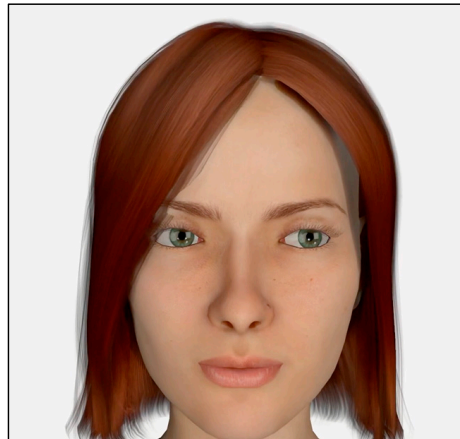**Figure A4.** Roxie looking to the participant's right while smiling.

Figure A5. Rex looking directly at the participant while nodding.

Figure A6. Rex looking to the participant's right while nodding.
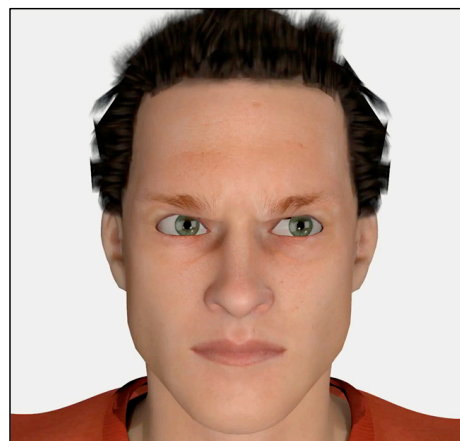

Figure A7. Roxie looking directly at the participant while nodding.

Figure A8. Roxie looking to the participant's right while nodding.
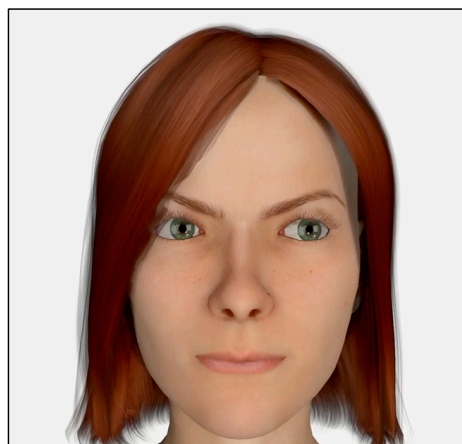

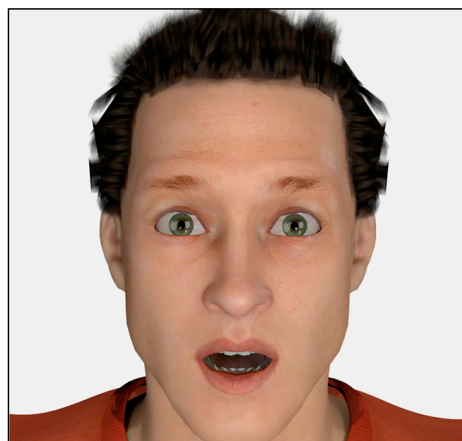**Figure A9.** Rex looking directly at the participant while making an angry expression.

Figure A10. Rex looking to the participant's right while making an angry expression.

Figure A11. Roxie looking directly at the participant while making an angry expression.

Figure A12. Roxie looking to the participant's right while making an angry expression.

Figure A13. Rex looking directly at the participant while making a surprised expression.

Figure A14. Rex looking to the participant's right while making a surprised expression.

Figure A15. Roxie looking directly at the participant while making a surprised expression.

Figure A16. Roxie looking to the participant's right while making a surprised expression.
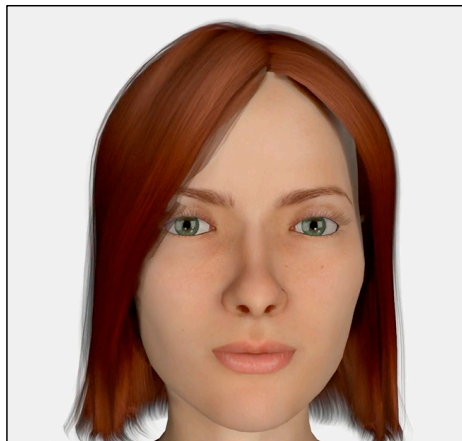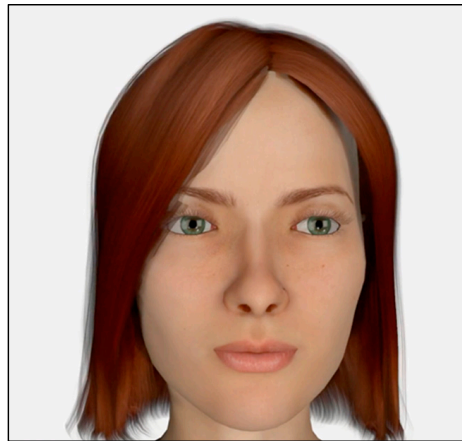
Figure A17. Rex looking directly at the participant while shaking its head.

Figure A18. Rex looking to the participant's right while shaking its head.
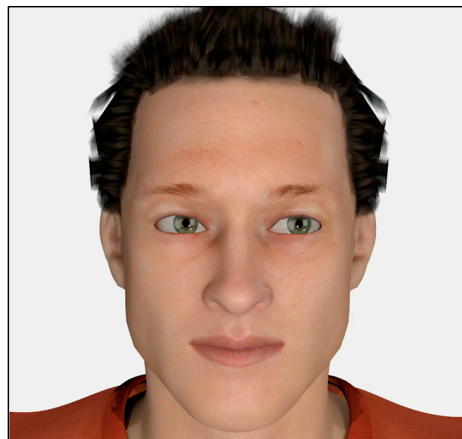

Figure A19. Roxie looking directly at the participant while shaking its head.

Figure A20. Roxie looking to the participant's right while shaking its head.

Figure A21. Rex looking directly at the participant with a neutral expression.

Figure A22. Rex looking to the participant's right with a neutral expression.

Figure A23. Roxie looking directly at the participant with a neutral expression.

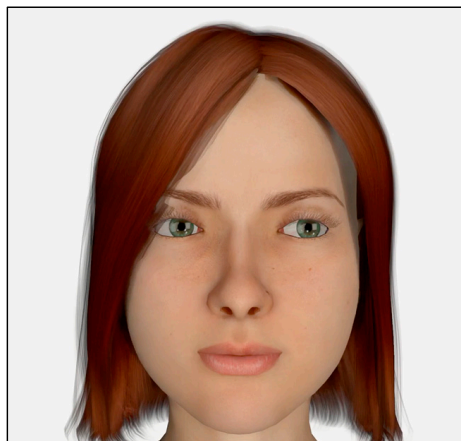**Figure A24.** Roxie looking to the participant's right with a neutral expression.

Figure A25. Rex looking directly at the participant while puffing its cheeks.

Figure A26. Rex looking to the participant's right while puffing its cheeks.

Figure A27. Roxie looking directly at the participant while puffing its cheeks.

Figure A28. Roxie looking to the participant's right while puffing its cheeks.

References

1. Rasouli, A.; Kotseruba, I.; Tsotsos, J.K. Understanding pedestrian behavior in complex traffic scenes. *IEEE Trans. Intell. Veh.* **2017**, *3*, 61–70. [CrossRef]
2. Markkula, G.; Madigan, R.; Nathanael, D.; Portouli, E.; Lee, Y.M.; Dietrich, A.; Billington, J.; Schieben, A.; Merat, N. Defining interactions: A conceptual framework for understanding interactive behaviour in human and automated road traffic. *Theor. Issues Ergon. Sci.* **2020**, *21*, 728–752. [CrossRef]
3. Färber, B. Communication and communication problems between autonomous vehicles and human drivers. In *Autonomous driving*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 125–144.

4. Sucha, M.; Dostal, D.; Risser, R. Pedestrian-driver communication and decision strategies at marked crossings. *Accid. Anal. Prev.* **2017**, *102*, 41–50. [CrossRef] [PubMed]
5. Llorca, D.F. From driving automation systems to autonomous vehicles: Clarifying the terminology. *arXiv* **2021**, arXiv:2103.10844.
6. SAE International. Taxonomy and Definitions of Terms Related to Driving Automation Systems for on-Road Motor Vehicles. 2021. Available online: www.sae.org (accessed on 26 January 2022).
7. Dey, D.; Terken, J. Pedestrian interaction with vehicles: Roles of explicit and implicit communication. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Oldenburg, Germany, 24–27 September 2017; pp. 109–113.
8. Moore, D.; Currano, R.; Strack, G.E.; Sirkin, D. 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, The Netherlands, 21–25 September 2019; pp. 295–307.
9. Lee, Y.M.; Madigan, R.; Giles, O.; Garach-Morcillo, L.; Markkula, G.; Fox, C.; Camara, F.; Rothmueller, M.; Vendelbo-Larsen, S.A.; Rasmussen, P.H.; et al. Road users rarely use explicit communication when interacting in today's traffic: Implications for automated vehicles. *Cogn. Technol. Work.* **2020**, *23*, 367–380. [CrossRef]
10. Guéguen, N.; Eyssartier, C.; Meineri, S. A pedestrian's smile and drivers' behavior: When a smile increases careful driving. *J. Saf. Res.* **2016**, *56*, 83–88. [CrossRef]
11. Guéguen, N.; Meineri, S.; Eyssartier, C. A pedestrian's stare and drivers' stopping behavior: A field experiment at the pedestrian crossing. *Saf. Sci.* **2015**, *75*, 87–89. [CrossRef]
12. Ren, Z.; Jiang, X.; Wang, W. Analysis of the influence of pedestrians' eye contact on drivers' comfort boundary during the crossing conflict. *Procedia Eng.* **2016**, *137*, 399–406. [CrossRef]
13. Nathanael, D.; Portouli, E.; Papakostopoulos, V.; Gkikas, K.; Amditis, A. Naturalistic observation of interactions between car drivers and pedestrians in high density urban settings. In *Congress of the International Ergonomics Association*; Springer: Berlin Germany, 2018; pp. 389–397.
14. Dey, D.; Walker, F.; Martens, M.; Terken, J. Gaze patterns in pedestrian interaction with vehicles: Towards effective design of external human-machine interfaces for automated vehicles. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Utrecht, The Netherlands, 21–25 September 2019; pp. 369–378.
15. Eisma, Y.B.; Van Bergen, S.; Ter Brake, S.M.; Hensen, M.T.T.; Tempelaar, W.J.; De Winter, J.C.F. External Human–Machine Interfaces: The Effect of Display Location on Crossing Intentions and Eye Movements. *Information* **2020**, *11*, 13. [CrossRef]
16. Uttley, J.; Lee, Y.M.; Madigan, R.; Merat, N. Road user interactions in a shared space setting: Priority and communication in a UK car park. *Transp. Res. Part F Traffic Psychol. Behav.* **2020**, *72*, 32–46. [CrossRef]
17. de Winter, J.; Bazilinskyy, P.; Wesdorp, D.; de Vlam, V.; Hopmans, B.; Visscher, J.; Dodou, D. How do pedestrians distribute their visual attention when walking through a parking garage? An eye-tracking study. *Ergonomics* **2021**, *64*, 793–805. [CrossRef]
18. Kong, X.; Das, S.; Zhang, Y.; Xiao, X. Lessons learned from pedestrian-driver communication and yielding patterns. *Transp. Res. Part F Traffic Psychol. Behav.* **2021**, *79*, 35–48. [CrossRef]
19. Onkhar, V.; Bazilinskyy, P.; Dodou, D.; De Winter, J.C.F. The effect of drivers' eye contact on pedestrians' perceived safety. *Transp. Res. Part F Traffic Psychol. Behav.* **2022**, *84*, 194–210. [CrossRef]
20. Lobjois, R.; Cavallo, V. Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task. *Accid. Anal. Prev.* **2007**, *39*, 934–943. [CrossRef]
21. Sun, R.; Zhuang, X.; Wu, C.; Zhao, G.; Zhang, K. The estimation of vehicle speed and stopping distance by pedestrians crossing streets in a naturalistic traffic environment. *Transp. Res. Part F Traffic Psychol. Behav.* **2015**, *30*, 97–106. [CrossRef]
22. Papić, Z.; Jović, A.; Simeunović, M.; Saulić, N.; Lazarević, M. Underestimation tendencies of vehicle speed by pedestrians when crossing unmarked roadway. *Accid. Anal. Prev.* **2020**, *143*, 105586. [CrossRef] [PubMed]
23. *ISO/TR 23049:2018*; Road Vehicles: Ergonomic aspects of external visual communication from automated vehicles to other road users. BSI: London, UK, 2018.
24. Merat, N.; Louw, T.; Madigan, R.; Wilbrink, M.; Schieben, A. What externally presented information do VRUs require when interacting with fully Automated Road Transport Systems in shared space? *Accid. Anal. Prev.* **2018**, *118*, 244–252. [CrossRef] [PubMed]
25. Rasouli, A.; Tsotsos, J.K. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 900–918. [CrossRef]
26. Rouchitsas, A.; Alm, H. External human–machine interfaces for autonomous vehicle-to-pedestrian communication: A review of empirical work. *Front. Psychol.* **2019**, *10*, 2757. [CrossRef]
27. Schieben, A.; Wilbrink, M.; Kettwich, C.; Madigan, R.; Louw, T.; Merat, N. Designing the interaction of automated vehicles with other traffic participants: Design considerations based on human needs and expectations. *Cogn. Technol. Work.* **2019**, *21*, 69–85. [CrossRef]
28. Carmona, J.; Guindel, C.; Garcia, F.; de la Escalera, A. eHMI: Review and Guidelines for Deployment on Autonomous Vehicles. *Sensors* **2021**, *21*, 2912. [CrossRef]
29. Ezzati Amini, R.; Katrakazas, C.; Riener, A.; Antoniou, C. Interaction of automated driving systems with pedestrians: Challenges, current solutions, and recommendations for eHMIs. *Transp. Rev.* **2021**, *41*, 788–813. [CrossRef]

30. Tabone, W.; de Winter, J.; Ackermann, C.; Bärgman, J.; Baumann, M.; Deb, S.; Stanton, N.A. Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. *Transp. Res. Interdisc. Perspect.* **2021**, *9*, 100293. [CrossRef]

31. Böckle, M.P.; Brenden, A.P.; Klingegård, M.; Habibovic, A.; Bout, M. SAV2P: Exploring the impact of an interface for shared automated vehicles on pedestrians' experience. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct, Oldenburg, Germany, 24–27 September 2017; pp. 136–140.

32. Chang, C.M.; Toda, K.; Sakamoto, D.; Igarashi, T. Eyes on a Car: An Interface Design for Communication between an Autonomous Car and a Pedestrian. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Oldenburg, Germany, 24–27 September 2017; pp. 65–73.

33. Costa, G. Designing Framework for Human-Autonomous Vehicle Interaction. Master's thesis, Keio University Graduate School of Media Design, Minato, Japan, 2017.

34. Deb, S.; Strawderman, L.J.; Carruth, D.W. Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment. *Transp. Res. Part F Traffic Psychol. Behav.* **2018**, *59*, 135–149. [CrossRef]

35. Habibovic, A.; Lundgren, V.M.; Andersson, J.; Klingegård, M.; Lagström, T.; Sirkka, A.; Fagerlönn, J.; Edgren, C.; Fredriksson, R.; Krupenia, S.; et al. Communicating intent of automated vehicles to pedestrians. *Front. Psychol.* **2018**, *9*, 1336. [CrossRef] [PubMed]

36. Hudson, C.R.; Deb, S.; Carruth, D.W.; McGinley, J.; Frey, D. Pedestrian perception of autonomous vehicles with external interacting features. In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Orlando, FL, USA, 22–26 July 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 33–39.

37. Mahadevan, K.; Somanath, S.; Sharlin, E. Communicating awareness and intent in autonomous vehicle-pedestrian interaction. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–12.

38. Othersen, I.; Conti-Kufner, A.S.; Dietrich, A.; Maruhn, P.; Bengler, K. Designing for automated vehicle and pedestrian communication: Perspectives on eHMIs from older and younger persons. *Proc. Hum. Factors Ergon. Soc. Eur.* **2018**, *4959*, 135–148.

39. Petzoldt, T.; Schleinitz, K.; Banse, R. Potential safety effects of a frontal brake light for motor vehicles. *IET Intell. Transp. Syst.* **2018**, *12*, 449–453. [CrossRef]

40. Song, Y.E.; Lehsing, C.; Fuest, T.; Bengler, K. External HMIs and their effect on the interaction between pedestrians and automated vehicles. In *International Conference on Intelligent Human Systems Integration, Dubai, United Arab Emirates, 7–9 January 2018*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 13–18.

41. de Clercq, K.; Dietrich, A.; Núñez Velasco, J.P.; De Winter, J.; Happee, R. External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. *Hum. Factors* **2019**, *61*, 1353–1370. [CrossRef]

42. Holländer, K.; Colley, A.; Mai, C.; Häkkilä, J.; Alt, F.; Pfleging, B. 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, 1–4 October 2019; pp. 1–11.

43. Stadler, S.; Cornet, H.; Novaes Theoto, T.; Frenkler, F. A tool, not a toy: Using virtual reality to evaluate the communication between autonomous vehicles and pedestrians. In *Augmented Reality and Virtual Reality*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 203–216.

44. Ackermans, S.C.A.; Dey, D.D.; Ruijten, P.A.; Cuijpers, R.H.; Pfleging, B. The effects of explicit intention communication, conspicuous sensors, and pedestrian attitude in interactions with automated vehicles. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Association for Computing Machinery, Inc.: New York, NY, USA, 2020; p. 70.

45. Faas, S.M.; Mathis, L.A.; Baumann, M. External HMI for self-driving vehicles: Which information shall be displayed? *Transp. Res. Part F Traffic Psychol. Behav.* **2020**, *68*, 171–186. [CrossRef]

46. Singer, T.; Kobbert, J.; Zandi, B.; Khanh, T.Q. Displaying the driving state of automated vehicles to other road users: An international, virtual reality-based study as a first step for the harmonized regulations of novel signaling devices. *IEEE Trans. Intell. Transp. Syst.* **2020**, *23*, 2904–2918. [CrossRef]

47. Lee, Y.M.; Madigan, R.; Uzondu, C.; Garcia, J.; Romano, R.; Markkula, G.; Merat, N. Learning to interpret novel eHMI: The effect of vehicle kinematics and eHMI familiarity on pedestrian' crossing behavior. *J. Saf. Res.* **2022**, *80*, 270–280. [CrossRef] [PubMed]

48. Wilbrink, M.; Lau, M.; Illgner, J.; Schieben, A.; Oehl, M. Impact of External Human–Machine Interface Communication Strategies of Automated Vehicles on Pedestrians' Crossing Decisions and Behaviors in an Urban Environment. *Sustainability* **2021**, *13*, 8396. [CrossRef]

49. Clamann, M.; Aubert, M.; Cummings, M.L. Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles. In Proceedings of the Transportation Research Board 96th Annual Meeting, Washington, DC, USA, 8–12 January 2017. No. 17-02119.

50. Li, Y.; Dikmen, M.; Hussein, T.G.; Wang, Y.; Burns, C. To cross or not to cross: Urgency-based external warning displays on autonomous vehicles to improve pedestrian crossing safety. In Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Toronto, ON, Canada, 23–25 September 2018; pp. 188–197.

51. Hensch, A.C.; Neumann, I.; Beggiato, M.; Halama, J.; Krems, J.F. How should automated vehicles communicate?—Effects of a light-based communication approach in a Wizard-of-Oz study. In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Washington, DC, USA, 24–28 July 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 79–91.

52. Dey, D.; Habibovic, A.; Löcken, A.; Wintersberger, P.; Pfleging, B.; Riener, A.; Martens, M.; Terken, J. Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transp. Res. Interdiscip. Perspect.* **2020**, *7*, 100174. [CrossRef]

53. Bevan, N.; Carter, J.; Harker, S. ISO 9241-11 revised: What have we learnt about usability since 1998? In Proceedings of the International Conference on Human-Computer Interaction, Bamberg, Germany, 14–18 September 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 143–151.

54. Fridman, L.; Mehler, B.; Xia, L.; Yang, Y.; Facusse, L.Y.; Reimer, B. To walk or not to walk: Crowdsourced assessment of external vehicle-to-pedestrian displays. *arXiv* **2017**, arXiv:1707.02698.

55. Ackermann, C.; Beggiato, M.; Schubert, S.; Krems, J.F. An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles? *Appl. Ergon.* **2019**, *75*, 272–282. [CrossRef] [PubMed]

56. Bazilinskyy, P.; Dodou, D.; De Winter, J. Survey on eHMI concepts: The effect of text, color, and perspective. *Transp. Res. Part F Traffic Psychol. Behav.* **2019**, *67*, 175–194. [CrossRef]

57. Eisma, Y.B.; Reiff, A.; Kooijman, L.; Dodou, D.; De Winter, J.C.F. External human-machine interfaces: Effects of message perspective. *Transp. Res. Part F Traffic Psychol. Behav.* **2021**, *78*, 30–41. [CrossRef]

58. Zhang, J.; Vinkhuyzen, E.; Cefkin, M. Evaluation of an autonomous vehicle external communication system concept: A survey study. In *International conference on applied human factors and ergonomics, Los Angeles, CA, USA, 17–21 July 2017*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 650–661.

59. Alvarez, W.M.; de Miguel, M.Á.; García, F.; Olaverri-Monreal, C. Response of Vulnerable Road Users to Visual Information from Autonomous Vehicles in Shared Spaces. In Proceedings of the 2019 IEEE, Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 27–30 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 3714–3719.

60. Chang, C.M. A Gender Study of Communication Interfaces between an Autonomous Car and a Pedestrian. In Proceedings of the 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Virtual Event, 21–22 September 2020; pp. 42–45.

61. Mirnig, N.; Perterer, N.; Stollnberger, G.; Tscheligi, M. Three strategies for autonomous car-to-pedestrian communication: A survival guide. In Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, Vienna, Austria, 6–9 March 2017; pp. 209–210.

62. Wickens, C.D.; Gordon, S.E.; Liu, Y.; Lee, J. *An Introduction to Human Factors Engineering*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2004; Volume 2.

63. Schilbach, L.; Wohlschlaeger, A.M.; Kraemer, N.C.; Newen, A.; Shah, N.J.; Fink, G.R.; Vogeley, K. Being with virtual others: Neural correlates of social interaction. *Neuropsychologia* **2006**, *44*, 718–730. [CrossRef]

64. Kuzmanovic, B.; Georgescu, A.L.; Eickhoff, S.B.; Shah, N.J.; Bente, G.; Fink, G.R.; Vogeley, K. Duration matters: Dissociating neural correlates of detection and evaluation of social gaze. *Neuroimage* **2009**, *46*, 1154–1163. [CrossRef] [PubMed]

65. Schrammel, F.; Pannasch, S.; Graupner, S.T.; Mojzisch, A.; Velichkovsky, B.M. Virtual friend or threat? The effects of facial expression and gaze interaction on psychophysiological responses and emotional experience. *Psychophysiology* **2009**, *46*, 922–931. [CrossRef] [PubMed]

66. Georgescu, A.L.; Kuzmanovic, B.; Schilbach, L.; Tepest, R.; Kulbida, R.; Bente, G.; Vogeley, K. Neural correlates of "social gaze" processing in high-functioning autism under systematic variation of gaze duration. *NeuroImage Clin.* **2013**, *3*, 340–351. [CrossRef] [PubMed]

67. Parsons, T.D. Virtual reality for enhanced ecological validity and experimental control in the clinical, affective, and social neurosciences. *Front. Hum. Neurosci.* **2015**, *9*, 660. [CrossRef]

68. Parsons, T.D.; Gaggioli, A.; Riva, G. Virtual reality for research in social neuroscience. *Brain Sci.* **2017**, *7*, 42. [CrossRef]

69. Dobs, K.; Bülthoff, I.; Schultz, J. Use and usefulness of dynamic face stimuli for face perception studies–a review of behavioral findings and methodology. *Front. Psychol.* **2018**, *9*, 1355. [CrossRef] [PubMed]

70. Georgescu, A.L.; Kuzmanovic, B.; Roth, D.; Bente, G.; Vogeley, K. The use of virtual characters to assess and train non-verbal communication in high-functioning autism. *Front. Hum. Neurosci.* **2014**, *8*, 807. [CrossRef] [PubMed]

71. Biocca, F.; Harms, C.; Burgoon, J.K. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence Teleoperators Virtual Environ.* **2003**, *12*, 456–480. [CrossRef]

72. Picard, R.W. *Affective Computing*; MIT Press: Cambridge, MA, USA, 2000.

73. Scherer, K.R.; Bänziger, T.; Roesch, E. (Eds.) *A Blueprint for Affective Computing: A Sourcebook and Manual*; Oxford University Press: Oxford, UK, 2010.

74. Cassell, J.; Thorisson, K.R. The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Appl. Artif. Intell.* **1999**, *13*, 519–538. [CrossRef]

75. Pütten, A.V.D.; Reipen, C.; Wiedmann, A.; Kopp, S.; Krämer, N.C. Comparing emotional vs. envelope feedback for ECAs. In Proceedings of the International Workshop on Intelligent Virtual Agents, Tokyo, Japan, 1–3 September 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 550–551.

76. Ochs, M.; Niewiadomski, R.; Pelachaud, C. How a virtual agent should smile? In Proceedings of the International Conference on Intelligent Virtual Agents, Philadelphia, PA, USA, 20–22 September 2010; Springer: Berlin/Heidelberg, Germany, 2010; pp. 427–440.

77. Wang, N.; Gratch, J. Don't just stare at me! In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 1241–1250.

78. McDonnell, R.; Breidt, M.; Bülthoff, H.H. Render me real?: Investigating the effect of render style on the perception of animated virtual humans. *ACM Trans. Graph. (TOG)* **2012**, *31*, 91. [CrossRef]

79. Wong, J.W.E.; McGee, K. Frown more, talk more: Effects of facial expressions in establishing conversational rapport with virtual agents. In Proceedings of the International Conference on Intelligent Virtual Agents, Santa Cruz CA, USA, 12–14 September 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 419–425.

80. Aljaroodi, H.M.; Adam, M.T.; Chiong, R.; Teubner, T. Avatars and embodied agents in experimental information systems research: A systematic review and conceptual framework. *Australas. J. Inf. Syst.* **2019**, *23*. [CrossRef]

81. Furuya, H.; Kim, K.; Bruder, G.; J Wisniewski, P.; Welch, F.G. Autonomous Vehicle Visual Embodiment for Pedestrian Interactions in Crossing Scenarios: Virtual Drivers in AVs for Pedestrian Crossing. In Proceedings of the Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; pp. 1–7.

82. Underwood, G. Visual attention and the transition from novice to advanced driver. *Ergonomics* **2007**, *50*, 1235–1249. [CrossRef] [PubMed]

83. Tafaj, E.; Kübler, T.C.; Kasneci, G.; Rosenstiel, W.; Bogdan, M. Online classification of eye tracking data for automated analysis of traffic hazard perception. In Proceedings of the International Conference on Artificial Neural Networks, Sofia, Bulgaria, 10–13 September 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 442–450.

84. Kaya, N.; Girgis, J.; Hansma, B.; Donmez, B. Hey, watch where you're going! An on-road study of driver scanning failures towards pedestrians and cyclists. *Accid. Anal. Prev.* **2021**, *162*, 106380. [CrossRef]

85. Von Grünau, M.; Anston, C. The detection of gaze direction: A stare-in-the-crowd effect. *Perception* **1995**, *24*, 1297–1313. [CrossRef] [PubMed]

86. Emery, N.J. The eyes have it: The neuroethology, function, and evolution of social gaze. *Neurosci. Biobehav. Rev.* **2000**, *24*, 581–604. [CrossRef]

87. Senju, A.; Hasegawa, T. Direct gaze captures visuospatial attention. *Vis. Cogn.* **2005**, *12*, 127–144. [CrossRef]

88. Senju, A.; Johnson, M.H. The eye contact effect: Mechanisms and development. *Trends Cogn. Sci.* **2009**, *13*, 127–134. [CrossRef] [PubMed]

89. Conty, L.; George, N.; Hietanen, J.K. Watching eyes effects: When others meet the self. *Conscious. Cogn.* **2016**, *45*, 184–197. [CrossRef] [PubMed]

90. Hamilton, A.F.D.C. Gazing at me: The importance of social meaning in understanding direct-gaze cues. *Philos. Trans. R. Soc. B Biol. Sci.* **2016**, *371*, 20150080. [CrossRef] [PubMed]

91. Frith, C.D.; Frith, U. Interacting minds–a biological basis. *Science* **1999**, *286*, 1692–1695. [CrossRef]

92. Gallagher, H.L.; Frith, C.D. Functional imaging of 'theory of mind'. *Trends Cogn. Sci.* **2003**, *7*, 77–83. [CrossRef]

93. Krumhuber, E.G.; Kappas, A.; Manstead, A.S. Effects of dynamic aspects of facial expressions: A review. *Emot. Rev.* **2013**, *5*, 41–46. [CrossRef]

94. Horstmann, G. What do facial expressions convey: Feeling states, behavioral intentions, or actions requests? *Emotion* **2003**, *3*, 150. [CrossRef]

95. Hess, U.; Adams, R.B., Jr.; Kleck, R.E. When Two Do the Same, It Might Not Mean the Same: The Perception of Emotional Expressions Shown by Men and Women. In *Group Dynamics and Emotional Expression*; Hess, U., Philippot, P., Eds.; Cambridge University Press: Cambridge, UK, 2007; pp. 33–50.

96. Scherer, K.R.; Grandjean, D. Facial expressions allow inference of both emotions and their components. *Cogn. Emot.* **2008**, *22*, 789–801. [CrossRef]

97. Ekman, P. Facial expressions of emotion: New findings, new questions. *Psychol. Sci.* **1992**, *3*, 34–38. [CrossRef]

98. Berkowitz, L.; Harmon-Jones, E. Toward an understanding of the determinants of anger. *Emotion* **2004**, *4*, 107. [CrossRef] [PubMed]

99. Schützwohl, A. Approach and avoidance during routine behavior and during surprise in a non-evaluative task: Surprise matters and so does the valence of the surprising event. *Front. Psychol.* **2018**, *9*, 826. [CrossRef]

100. Reisenzein, R.; Horstmann, G.; Schützwohl, A. The cognitive-evolutionary model of surprise: A review of the evidence. *Top. Cogn. Sci.* **2019**, *11*, 50–74. [CrossRef]

101. Nusseck, M.; Cunningham, D.W.; Wallraven, C.; Bülthoff, H.H. The contribution of different facial regions to the recognition of conversational expressions. *J. Vis.* **2008**, *8*, 1. [CrossRef] [PubMed]

102. Cunningham, D.W.; Wallraven, C. Dynamic information for the recognition of conversational expressions. *J. Vis.* **2009**, *9*, 7. [CrossRef] [PubMed]

103. Kaulard, K.; Cunningham, D.W.; Bülthoff, H.H.; Wallraven, C. The MPI facial expression database—A validated database of emotional and conversational facial expressions. *PLoS ONE* **2012**, *7*, e32321. [CrossRef]

104. Kendon, A. Some uses of the head shake. *Gesture* **2002**, *2*, 147–182. [CrossRef]

105. Guidetti, M. Yes or no? How young French children combine gestures and speech to agree and refuse. *J. Child Lang.* **2005**, *32*, 911–924. [CrossRef] [PubMed]

106. Andonova, E.; Taylor, H.A. Nodding in dis/agreement: A tale of two cultures. *Cogn. Processing* **2012**, *13*, 79–82. [CrossRef] [PubMed]

107. Fusaro, M.; Vallotton, C.D.; Harris, P.L. Beside the point: Mothers' head nodding and shaking gestures during parent–child play. *Infant Behav. Dev.* **2014**, *37*, 235–247. [CrossRef] [PubMed]

108. Osugi, T.; Kawahara, J.I. Effects of Head Nodding and Shaking Motions on Perceptions of Likeability and Approachability. *Perception* **2018**, *47*, 16–29. [CrossRef]

109. Moretti, S.; Greco, A. Nodding and shaking of the head as simulated approach and avoidance responses. *Acta Psychol.* **2020**, *203*, 102988. [CrossRef]

110. Semcon. The Smiling Car. 2016. Available online: https://semcon.com/uk/smilingcar/ (accessed on 21 April 2022).

111. Becker, D.V.; Kenrick, D.T.; Neuberg, S.L.; Blackwell, K.C.; Smith, D.M. The confounded nature of angry men and happy women. *J. Personal. Soc. Psychol.* **2007**, *92*, 179. [CrossRef]

112. Niedenthal, P.M.; Mermillod, M.; Maringer, M.; Hess, U. The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behav. Brain Sci.* **2010**, *33*, 417. [CrossRef] [PubMed]

113. Barrett, L.F.; Adolphs, R.; Marsella, S.; Martinez, A.M.; Pollak, S.D. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychol. Sci. Public Interest* **2019**, *20*, 1–68. [CrossRef] [PubMed]

114. Weber, M.; Giacomin, J.; Malizia, A.; Skrypchuk, L.; Gkatzidou, V.; Mouzakitis, A. Investigation of the dependency of the drivers' emotional experience on different road types and driving conditions. *Transp. Res. Part F Traffic Psychol. Behav.* **2019**, *65*, 107–120. [CrossRef]

115. Popuşoi, S.A.; Havârneanu, G.M.; Havârneanu, C.E. "Get the f#∗k out of my way!" Exploring the cathartic effect of swear words in coping with driving anger. *Transp. Res. Part F Traffic Psychol. Behav.* **2018**, *56*, 215–226.

116. Stephens, A.N.; Lennon, A.; Bihler, C.; Trawley, S. The measure for angry drivers (MAD). *Transp. Res. Part F Traffic Psychol. Behav.* **2019**, *64*, 472–484. [CrossRef]

117. Deffenbacher, J.L.; Lynch, R.S.; Oetting, E.R.; Swaim, R.C. The Driving Anger Expression Inventory: A measure of how people express their anger on the road. *Behav. Res. Ther.* **2002**, *40*, 717–737. [CrossRef]

118. Kim, H.; Somerville, L.H.; Johnstone, T.; Alexander, A.L.; Whalen, P.J. Inverse amygdala and medial prefrontal cortex responses to surprised faces. *Neuroreport* **2003**, *14*, 2317–2322. [CrossRef] [PubMed]

119. Marsh, A.A.; Ambady, N.; Kleck, R.E. The effects of fear and anger facial expressions on approach-and avoidance-related behaviors. *Emotion* **2005**, *5*, 119. [CrossRef]

120. Neta, M.; Davis, F.C.; Whalen, P.J. Valence resolution of ambiguous facial expressions using an emotional oddball task. *Emotion* **2011**, *11*, 1425. [CrossRef] [PubMed]

121. Neta, M.; Whalen, P.J. The primacy of negative interpretations when resolving the valence of ambiguous facial expressions. *Psychol. Sci.* **2010**, *21*, 901–907. [CrossRef] [PubMed]

122. Yamada, H.; Matsuda, T.; Watari, C.; Suenaga, T. Dimensions of visual information for categorizing facial expressions of emotion. *Jpn. Psychol. Res.* **1994**, *35*, 172–181. [CrossRef]

123. Tottenham, N.; Tanaka, J.W.; Leon, A.C.; McCarry, T.; Nurse, M.; Hare, T.A.; Nelson, C. The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Res.* **2009**, *168*, 242–249. [CrossRef]

124. Wu, S.; Sun, S.; Camilleri, J.A.; Eickhoff, S.B.; Yu, R. Better the devil you know than the devil you don't: Neural processing of risk and ambiguity. *NeuroImage* **2021**, *236*, 118109. [CrossRef]

125. Alter, A.L.; Oppenheimer, D.M.; Epley, N.; Eyre, R.N. Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *J. Exp. Psychol. Gen.* **2007**, *136*, 569. [CrossRef] [PubMed]

126. Naples, A.; Nguyen-Phuc, A.; Coffman, M.; Kresse, A.; Faja, S.; Bernier, R.; McPartland, J.C. A computer-generated animated face stimulus set for psychophysiological research. *Behav. Res. Methods* **2015**, *47*, 562–570. [CrossRef]

127. Nelson, N.L.; Russell, J.A. Building emotion categories: Children use a process of elimination when they encounter novel expressions. *J. Exp. Child Psychol.* **2016**, *151*, 120–130. [CrossRef]

128. Wiese, E.; Metta, G.; Wykowska, A. Robots as intentional agents: Using neuroscientific methods to make robots appear more social. *Front. Psychol.* **2017**, *8*, 1663. [CrossRef] [PubMed]

129. Gamer, M.; Hecht, H. Are you looking at me? Measuring the cone of gaze. *J. Exp. Psychol. Hum. Percept. Perform.* **2007**, *33*, 705. [CrossRef]

130. Onkhar, V.; Bazilinskyy, P.; Stapel, J.C.J.; Dodou, D.; Gavrila, D.; De Winter, J.C.F. Towards the detection of driver–pedestrian eye contact. *Pervasive Mob. Comput.* **2021**, *76*, 101455. [CrossRef]

131. Kohler, C.G.; Turner, T.; Stolar, N.M.; Bilker, W.B.; Brensinger, C.M.; Gur, R.E.; Gur, R.C. Differences in facial expressions of four universal emotions. *Psychiatry Res.* **2004**, *128*, 235–244. [CrossRef]

132. Ambadar, Z.; Cohn, J.F.; Reed, L.I. All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *J. Nonverbal Behav.* **2009**, *33*, 17–34. [CrossRef] [PubMed]

133. Helwig, N.E.; Sohre, N.E.; Ruprecht, M.R.; Guy, S.J.; Lyford-Pike, S. Dynamic properties of successful smiles. *PLoS ONE* **2017**, *12*, e0179708. [CrossRef] [PubMed]

134. N'Diaye, K.; Sander, D.; Vuilleumier, P. Self-relevance processing in the human amygdala: Gaze direction, facial expression, and emotion intensity. *Emotion* **2009**, *9*, 798. [CrossRef]

135. Bantoft, C.; Summers, M.J.; Tranent, P.J.; Palmer, M.A.; Cooley, P.D.; Pedersen, S.J. Effect of standing or walking at a workstation on cognitive function: A randomized counterbalanced trial. *Hum. Factors* **2016**, *58*, 140–149. [CrossRef] [PubMed]

136. Kang, S.H.; Lee, J.; Jin, S. Effect of standing desk use on cognitive performance and physical workload while engaged with high cognitive demand tasks. *Appl. Ergon.* **2021**, *92*, 103306. [CrossRef] [PubMed]

137. Kaß, C.; Schoch, S.; Naujoks, F.; Hergeth, S.; Keinath, A.; Neukum, A. Standardized Test Procedure for External Human–Machine Interfaces of Automated Vehicles. *Information* **2020**, *11*, 173. [CrossRef]

138. Field, A. Discovering statistics using IBM SPSS statistics. Sage: Los Angeles, CA, USA,, 2013.

139. Smith, M.L.; Cottrell, G.W.; Gosselin, F.; Schyns, P.G. Transmitting and decoding facial expressions. *Psychol. Sci.* **2005**, *16*, 184–189. [CrossRef] [PubMed]

140. Summers, R.J.; Meese, T.S. The influence of fixation points on contrast detection and discrimination of patches of grating: Masking and facilitation. *Vis. Res.* **2009**, *49*, 1894–1900. [CrossRef] [PubMed]

141. Reisberg, D. *Cognition: Exploring the Science of the Mind: Sixth International Student Edition*; WW Norton & Company: New York, NY, USA, 2015.

142. Richler, J.J.; Mack, M.L.; Gauthier, I.; Palmeri, T.J. Holistic processing of faces happens at a glance. *Vis. Res.* **2009**, *49*, 2856–2861. [CrossRef] [PubMed]

143. Hershler, O.; Hochstein, S. At first sight: A high-level pop out effect for faces. *Vis. Res.* **2005**, *45*, 1707–1724. [CrossRef]

144. Jing, P.; Xu, G.; Chen, Y.; Shi, Y.; Zhan, F. The determinants behind the acceptance of autonomous vehicles: A systematic review. *Sustainability* **2020**, *12*, 1719. [CrossRef]

145. Tapiro, H.; Oron-Gilad, T.; Parmet, Y. Pedestrian distraction: The effects of road environment complexity and age on pedestrian's visual attention and crossing behavior. *J. Saf. Res.* **2020**, *72*, 101–109. [CrossRef]

146. Bainbridge, L. Ironies of automation. In *Analysis, Design and Evaluation of Man–Machine Systems*; Pergamon Press: Oxford, UK, 1983; pp. 129–135.

147. Reason, J. Understanding adverse events: Human factors. *BMJ Qual. Saf.* **1995**, *4*, 80–89. [CrossRef]

148. Colley, M.; Walch, M.; Rukzio, E. Unveiling the Lack of Scalability in Research on External Communication of Autonomous Vehicles. In Proceedings of the Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–9.

149. Faria, J.J.; Krause, S.; Krause, J. Collective behavior in road crossing pedestrians: The role of social information. *Behav. Ecol.* **2010**, *21*, 1236–1242. [CrossRef]

150. Lanzer, M.; Baumann, M. Does crossing the road in a group influence pedestrians' gaze behavior? *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2020**, *64*, 1938–1942. [CrossRef]

151. Wilbrink, M.; Nuttelmann, M.; Oehl, M. Scaling up Automated Vehicles' eHMI Communication Designs to Interactions with Multiple Pedestrians–Putting eHMIs to the Test. In Proceedings of the 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Leeds, UK, 9–10 September 2021; pp. 119–122.

152. Dey, D.; van Vastenhoven, A.; Cuijpers, R.H.; Martens, M.; Pfleging, B. Towards Scalable eHMIs: Designing for AV-VRU Communication Beyond One Pedestrian. In Proceedings of the 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Leeds, UK, 9–14 September 2021; pp. 274–286.

153. Joisten, P.; Liu, Z.; Theobald, N.; Webler, A.; Abendroth, B. Communication of Automated Vehicles and Pedestrian Groups: An Intercultural Study on Pedestrians' Street Crossing Decisions. In Proceedings of the Mensch und Computer 2021-Tagungsband, Ingolstadt, Germany, 5–8 September 2021.

154. Mayer, R.C.; Davis, J.H.; Schoorman, F.D. An integrative model of organizational trust. *Acad. Manag. Rev.* **1995**, *20*, 709–734. [CrossRef]

155. Hoff, K.A.; Bashir, M. Trust in automation: Integrating empirical evidence on factors that influence trust. *Hum. Factors* **2015**, *57*, 407–434. [CrossRef]

156. Lee, J.D.; See, K.A. Trust in automation: Designing for appropriate reliance. *Hum. Factors* **2004**, *46*, 50–80. [CrossRef]

157. Dixon, L. Autonowashing: The greenwashing of vehicle automation. *Transp. Res. Interdiscip. Perspect.* **2020**, *5*, 100113. [CrossRef]

158. Faas, S.M.; Kraus, J.; Schoenhals, A.; Baumann, M. 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, 8–13 May 2021; pp. 1–17.

159. Choi, J.K.; Ji, Y.G. Investigating the importance of trust on adopting an autonomous vehicle. *Int. J. Hum.-Comput. Interact.* **2015**, *31*, 692–702. [CrossRef]

160. Hengstler, M.; Enkel, E.; Duelli, S. Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices. *Technol. Forecast. Soc. Change* **2016**, *105*, 105–120. [CrossRef]

161. Reig, S.; Norman, S.; Morales, C.G.; Das, S.; Steinfeld, A.; Forlizzi, J. A field study of pedestrians and autonomous vehicles. In Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications, Toronto, ON, Canada, 23–25 September 2018; pp. 198–209.

162. Oliveira, L.; Proctor, K.; Burns, C.G.; Birrell, S. Driving style: How should an automated vehicle behave? *Information* **2019**, *10*, 219. [CrossRef]

163. Olaverri-Monreal, C. Promoting trust in self-driving vehicles. *Nat. Electron.* **2020**, *3*, 292–294. [CrossRef]

164. Wang, Y.; Hespanhol, L.; Tomitsch, M. How Can Autonomous Vehicles Convey Emotions to Pedestrians? A Review of Emotionally Expressive Non-Humanoid Robots. *Multimodal Technol. Interact.* **2021**, *5*, 84. [CrossRef]

165. Nowak, K.L.; Rauh, C. Choose your "buddy icon" carefully: The influence of avatar androgyny, anthropomorphism, and credibility in online interactions. *Comput. Hum. Behav.* **2008**, *24*, 1473–1493. [CrossRef]

166. de Visser, E.J.; Krueger, F.; McKnight, P.; Scheid, S.; Smith, M.; Chalk, S.; Parasuraman, R. The world is not enough: Trust in cognitive agents. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2012**, *56*, 263–267. [CrossRef]

167. Pak, R.; Fink, N.; Price, M.; Bass, B.; Sturre, L. Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. *Ergonomics* **2012**, *55*, 1059–1072. [CrossRef]

168. Waytz, A.; Heafner, J.; Epley, N. The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *J. Exp. Soc. Psychol.* **2014**, *52*, 113–117. [CrossRef]

169. Kaleefathullah, A.A.; Merat, N.; Lee, Y.M.; Eisma, Y.B.; Madigan, R.; Garcia, J.; Winter, J.D. External Human–Machine Interfaces Can Be Misleading: An Examination of Trust Development and Misuse in a CAVE-Based Pedestrian Simulation Environment. *Hum. Factors* **2020**, *64*, 1070–1085. [CrossRef] [PubMed]

170. Matthews, M.; Chowdhary, G.; Kieson, E. Intent communication between autonomous vehicles and pedestrians. *arXiv* **2017**, arXiv:1708.07123.

171. Schweitzer, M.E.; Hershey, J.C.; Bradlow, E.T. Promises and lies: Restoring violated trust. *Organ. Behav. Hum. Decis. Processes* **2006**, *101*, 1–19. [CrossRef]

172. Holländer, K.; Wintersberger, P.; Butz, A. Overtrust in external cues of automated vehicles: An experimental investigation. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Utrecht, The Netherlands, 21–25 September 2019; pp. 211–221.

173. Gong, L. How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Comput. Hum. Behav.* **2008**, *24*, 1494–1509. [CrossRef]

174. Andrade, C. Internal, external, and ecological validity in research design, conduct, and evaluation. *Indian J. Psychol. Med.* **2018**, *40*, 498–499. [CrossRef] [PubMed]

175. Feldstein, I.; Dietrich, A.; Milinkovic, S.; Bengler, K. A pedestrian simulator for urban crossing scenarios. *IFAC-Pap.* **2016**, *49*, 239–244. [CrossRef]

176. Deb, S.; Carruth, D.W.; Sween, R.; Strawderman, L.; Garrison, T.M. Efficacy of virtual reality in pedestrian safety research. *Appl. Ergon.* **2017**, *65*, 449–460. [CrossRef]

177. Hassin, R.R.; Aviezer, H.; Bentin, S. Inherently ambiguous: Facial expressions of emotions, in context. *Emot. Rev.* **2013**, *5*, 60–65. [CrossRef]

178. Eisele, D.; Petzoldt, T. Effects of traffic context on eHMI icon comprehension. *Transp. Res. Part F Traffic Psychol. Behav.* **2022**, *85*, 1–12. [CrossRef]

179. Cavallo, V.; Dommes, A.; Dang, N.T.; Vienne, F. A street-crossing simulator for studying and training pedestrians. *Transp. Res. Part F Traffic Psychol. Behav.* **2019**, *61*, 217–228. [CrossRef]

180. Faas, S.M.; Mattes, S.; Kao, A.C.; Baumann, M. Efficient Paradigm to Measure Street-Crossing Onset Time of Pedestrians in Video-Based Interactions with Vehicles. *Information* **2020**, *11*, 360. [CrossRef]

181. Vermersch, P. Describing the practice of introspection. *J. Conscious. Stud.* **2009**, *16*, 20–57.

182. Cahour, B.; Salembier, P.; Zouinar, M. Analyzing lived experience of activity. *Le Trav. Hum.* **2016**, *79*, 259–284. [CrossRef]

183. Utriainen, R.; Pöllänen, M. Prioritizing Safety or Traffic Flow? Qualitative Study on Highly Automated Vehicles' Potential to Prevent Pedestrian Crashes with Two Different Ambitions. *Sustainability* **2020**, *12*, 3206. [CrossRef]

184. Deb, S.; Carruth, D.W.; Fuad, M.; Stanley, L.M.; Frey, D. Comparison of Child and Adult Pedestrian Perspectives of External Features on Autonomous Vehicles Using Virtual Reality Experiment. In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Washington, DC, USA, 24–28 July 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 145–156.

185. Tapiro, H.; Meir, A.; Parmet, Y.; Oron-Gilad, T. Visual search strategies of child-pedestrians in road crossing tasks. *Proc. Hum. Factors Ergon. Soc. Eur.* **2014**, 119–130.

186. Charisi, V.; Habibovic, A.; Andersson, J.; Li, J.; Evers, V. Children's views on identification and intention communication of self-driving vehicles. In Proceedings of the 2017 Conference on Interaction Design and Children, Stanford, CA, USA, 27–30 June 2017; pp. 399–404.

187. Klin, A.; Jones, W.; Schultz, R.; Volkmar, F.; Cohen, D. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch. Gen. Psychiatry* **2002**, *59*, 809–816. [CrossRef] [PubMed]

188. Crehan, E.T.; Althoff, R.R. Me looking at you, looking at me: The stare-in-the-crowd effect and autism spectrum disorder. *J. Psychiatr. Res.* **2021**, *140*, 101–109. [CrossRef] [PubMed]

189. Strauss, D.; Shavelle, R.; Anderson, T.W.; Baumeister, A. External causes of death among persons with developmental disability: The effect of residential placement. *Am. J. Epidemiol.* **1998**, *147*, 855–862. [CrossRef] [PubMed]