External human–machine interfaces
Gimmick or necessity?
de Winter, Joost; Dodou, Dimitra

DOI
10.1016/j.trip.2022.100643

Publication date
2022

Document Version
Final published version

Published in
Transportation Research Interdisciplinary Perspectives

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.
External human–machine interfaces: Gimmick or necessity?

Joost de Winter*, Dimitra Dodou

Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, the Netherlands

1. Introduction

The last few years have seen a surge of interest in automated driving. This trend can be identified in academia, signified by a large number of publications (Ayoub et al., 2019; Gandia et al., 2019), and in industry, signified by numerous news items and forecasts (Deloitte, 2019; Litman, 2021). This surge of interest may be a genuine reflection of advancements in technology, including progress in sensor systems, computational speed, and computer vision. However, there is still much discussion about whether eHMIs are needed. The present article surveys arguments for and against eHMIs. We list three arguments against eHMIs: (1) Implicit communication dominates pedestrian-AV interaction, and there is no social interaction void to be filled, (2) There is a large variety of eHMI concepts and a lack of standardization and consensus, and (3) eHMIs may elicit various negative effects such as distraction, confusion, and overreliance. Next, we present five reasons why eHMIs may be useful or required: (1) eHMIs can make planned actions of the AV visible, thereby increasing the efficiency of pedestrian-AV interaction, (2) Participants value an eHMI compared to no eHMI, (3) eHMIs do not have to be limited to showing instructions or the AV’s planned actions; showing the AV’s mode or the AV’s cooperative or detection capabilities are other uses of eHMIs, (4) Recent research shows that driver eye contact is important in traffic, and a social interaction void thus exists, and (5) A large portion of pedestrian-vehicle accidents in current traffic is caused by unclear implicit communication, suggesting that pedestrians may benefit from explicit eHMIs. It is hoped that this article contributes to the critical discussion of whether eHMIs are needed and how they should be designed.

* Corresponding author.
E-mail addresses: j.c.f.dewinter@tudelft.nl (J. de Winter), d.dodou@tudelft.nl (D. Dodou).

https://doi.org/10.1016/j.trip.2022.100643
Received 18 February 2022; Received in revised form 5 June 2022; Accepted 13 June 2022
Available online 4 July 2022

© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
end with closing statements.

2. Arguments against eHMIs

2.1. Argument 1 against eHMIs: Implicit communication dominates; no social interaction void exists

Several scientists argue that research in AV-pedestrian interaction should focus on how the AV behaves (also called: implicit communication) and that one should not add bells and whistles in the form of eHMIs. Their line of reasoning is that in current vehicle–pedestrian interaction, explicit modes of communication, such as eye contact, hand gestures, nodding, or using the high beams, are rare (e.g., Lee et al., 2021) and that the above-mentioned social interaction void is therefore trivial or nonexistent. That is, pedestrians are able to cross the road without using explicit communication, and implicit communication is the primary cue that pedestrians use (Moore et al., 2019).

More specifically, eHMIs are claimed to substitute driver eye contact. However, in current traffic, pedestrians often cannot even see the driver, let alone establish eye contact or perceive other gestures. Driving at night or a driver wearing sunglasses, windshield glare, sunlight, and shadows are some of the reasons why the driver’s eyes may not be visible (AlAdawy et al., 2019). Moore et al. (2019) showed that pedestrians usually cross the road without noticing the AV’s features and hardly adjust their walking behavior, even when there is no driver in the driver’s seat at all. Besides, in current traffic, there appears to be no need to employ eHMIs (other than the already existing cueing systems such as the horn, blinkers, brake lights, and headlamps), so why then would we need to deploy eHMIs on AVs?

Other authors do not outright reject the need for eHMIs but portray eHMIs as of secondary importance. For example, Domeyer et al. (2020) emphasized that implicit communication has precedence over explicit communication: “the concepts [of motion-based communication] described here do not invalidate the need for research on lighting and other explicit signals, it suggests that behavior is the basis for on-road communication”. Similarly, Lee et al. (2021) concluded that “road users rarely used explicit communication to convey information about crossing intentions” (p. 377), and they implied “that there may be limited requirement for automated vehicles to adopt explicit communication solutions ...” (p. 378).

2.2. Argument 2 against eHMIs: There is a lack of standardization and consensus

A large variety of eHMI concepts have been proposed so far. Interest in eHMIs arose around 2015 and 2016 when vehicle manufacturers introduced futuristic-looking concept cars that could communicate via eHMIs. LED strips, text messages, and projections on the ground were among the most eye-catching designs (for a review of 22 eHMI concepts from the industry, see Bazilinskyy et al., 2019). Academia enthusiastically picked up the topic of eHMIs and continued the investigation. A survey of the literature by Dey et al. (2020a) identified 70 concepts available by mid-2019, a number that continues to grow. eHMIs take many forms, including eHMIs that speak (Mahadevan et al., 2018), googly eyes (Chang et al., 2017), a smile (Deb et al., 2018), and laser-like displays (Dietrich et al., 2018; Mok et al., 2022). Dey et al. (2020a) referred to the present situation as an “eHMI jungle”.

When reading Dey et al.’s review paper, one cannot avoid the impression that the field lacks consensus and standardization. Current attempts at standardization of eHMIs (ISO 23049:2018; International Organization for Standardization, 2018) appear to be of preliminary and suggestive nature: The ISO document provides no recommendations about the visual appearance of the communication, other than that “care should be taken such that consistency or coherency with existing vehicle interfaces is maintained” (p. 4). While there is research concerned with technical specifications of eHMIs, such as display size, luminance, and the distinguishability and interpretation of colors and animations (Blankenbach et al., 2022; Clämann et al., 2017; GRE Autonomous Vehicle Signalling Requirements, 2019; Werner, 2018), concrete recommendations regarding eHMI design seem lacking.

An inspection of the literature shows there is disagreement about some of the fundamentals of eHMI design. Human Factors experts (Tabone et al., 2021a) and current standards (International Organization for Standardization, 2018) suggest that an eHMI should not instruct, such as via the text WALK. The reasoning behind this is that an instructive message can cause accidents if a non-automated vehicle is arriving simultaneously or if the message is picked up by a pedestrian for whom the message was not intended. At the same time, it can be argued that an instructive eHMI is unambiguous and safe to use if the AV can ascertain that it is indeed safe to cross at that moment, just like current (pedestrian) traffic lights are used. Researchers especially caution against text-based eHMIs, because text requires focused attention (Cefkin, 2018; Dey et al., 2022). However, an advantage of text-based eHMIs is that they can be understood directly (language barriers not considered), while non-textual eHMIs require training or experience (Bazilinskyy et al., 2019; De Clercq et al., 2019).

Indeed, evaluations of eHMIs show that instructive text-based messages such as WALK or DON’T WALK are relatively unambiguous and processed efficiently by users (Bazilinskyy et al., 2019, 2022; Ferenchak and Shafique, 2022; Guo et al., 2022). Currently, there appears to be no clarity in the literature about whether (instructive) text messages should or should not be used.

Another design choice of importance concerns the color of the eHMI. Cyan is an often recommended color because of its neutrality (as opposed to red or green). However, research shows that, depending on its precise tone, cyan risks confusion with green (Bazilinskyy et al., 2020; Dey et al., 2020b; GRE Autonomous Vehicle Signalling Requirements, 2019). It is not currently clear which colors are recommendable, whether it be cyan, yellow, orange, or purple.

Finally, it is unknown how eHMIs should be deployed in actual traffic. As it turns out, most of the eHMI research to date has been conducted in default scenarios, where a single pedestrian wants to cross the road in front of a single eHMI-equipped AV. However, it is easily possible to imagine a situation where two or more AVs give cues to a pedestrian, a situation that does not seem to have been considered so far. Future AVs may have to have to rely on connectivity and collaborative perception (Chen et al., 2019) to be able to understand the intentions of other road users. In turn, these developments raise the question of which agent in future connected traffic (e.g., AV, cloud) will have the authority to advise or instruct other agents. How eHMIs should address a specific pedestrian when multiple road users are present is another understudied topic (Colley et al., 2020), which has been addressed by several researchers but not resolved (Dey et al., 2021b; Dietrich et al., 2018; Hübner et al., 2022; Joisten et al., 2021; Verstegen et al., 2021; Wilbrink et al., 2021). A possible solution would be to provide personalized information to the pedestrian, such as via augmented reality (Hasan and Hasan, 2022; Tabone et al., 2021b; Tran et al., 2022).

1 The ISO document also provides little guidance on the type of modality to use. It first mentions that “… visual signalling is recommended”, while later in the document leaves the choice of modality open by highlighting the advantages and disadvantages of auditory versus visual signaling. The ISO document also states that signals should be “distinct and salient yet not distractive”, but does not specify how this could be accomplished.
Other types of vulnerable road users, such as cyclists, have not been considered much in eHMI research (exceptions are Bazilinskyy et al., 2021; Berge et al., 2022; GRE Autonomous Vehicle Signalling Requirements, 2019; Hou et al., 2020; Kaß et al., 2020a; b; Kunst et al., 2022; Li et al., 2021; Verstegen et al., 2021; Vlakveld et al., 2020; Yang et al., in press; and see Von Sawitzky et al., 2020, for Augmented Reality concepts for cyclists). Cyclists may interact with AVs in more dynamic situations, and may necessitate eHMIs that can be viewed from all sides (e.g., Vlakveld et al., 2020). Additionally, researchers have devised eHMIs that communicate to manual vehicles, such as in a bottleneck scenario (Rettenmaier et al., 2020) or to communicate yielding intention at junctions (e.g., Avsar et al., 2021; Mirmig et al., 2021; Papakostopoulos et al., 2021). These developments raise the question of whether eHMIs will have to be separately tailored to different road users (pedestrians, cyclists, manual vehicles) and scenarios.

At the moment, it is not clear how the above challenges should be addressed and how to blend eHMIs with existing signals such as traffic lights, turn indicators, and the horn. It is conceivable that the above scenarios could be solved. Researchers have devised eHMIs that communicate to manual vehicles, such as in a bottleneck scenario (Rettenmaier et al., 2020) or to communicate yielding intention at junctions (e.g., Avsar et al., 2021; Mirmig et al., 2021; Papakostopoulos et al., 2021). These developments raise the question of whether eHMIs will have to be separately tailored to different road users (pedestrians, cyclists, manual vehicles) and scenarios.

At the moment, it is not clear how the above challenges should be addressed and how to blend eHMIs with existing signals such as traffic lights, turn indicators, and the horn. It is conceivable that the above scenarios could be solved. Researchers have devised eHMIs that communicate to manual vehicles, such as in a bottleneck scenario (Rettenmaier et al., 2020) or to communicate yielding intention at junctions (e.g., Avsar et al., 2021; Mirmig et al., 2021; Papakostopoulos et al., 2021). These developments raise the question of whether eHMIs will have to be separately tailored to different road users (pedestrians, cyclists, manual vehicles) and scenarios.

2.3. Argument 3 against eHMIs: eHMIs elicit negative effects

It has often been argued that current traffic is visually demanding and that distractions are a significant risk for drivers (Horberry et al., 2006) and pedestrians (Tapiro et al., 2020). It can be argued that adding light sources in the form of eHMIs will exacerbate this problem. Norman and Emenegger noted: “If messages are placed outside of the car, human attention becomes more scattered, especially when there are many vehicles, so they may miss a critical signal” (as quoted in Tabone et al., 2021a; p. 10).

Most of the eHMI research thus far has been conducted online or in virtual-reality environments, with little opportunity for visual or cognitive distraction. There is a limited but growing number of eHMI studies conducted with real vehicles, but typically in simple settings such as parking lots (Ahn et al., 2021; Chen et al., 2020; Hensch et al., 2020; Liu et al., 2021), indoor environments (Burns et al., 2019; Reschke et al., 2018), test tracks (Faas et al., 2017; Fuest et al., 2020; Horn et al., 2021), or roads with otherwise restricted access (Barendse, 2019; Dey et al., 2021a; Habibovic et al., 2018; Joisten et al., 2019; Morales Alvarez et al., 2019; Mühmann, 2019; Papakostopoulos et al., 2021; Zadeh Darrehshourian, 2021). Research in real traffic is still relatively rare (Cefkin et al., 2019; Forke et al., 2021; Merat et al., 2018; Mirmig et al., 2021; Monzel et al., 2021), and some evidence concurs that eHMIs will have to compete with other visual cues in the environment. In particular, Cefkin et al. (2019) found that, in real traffic, which can be busy and requires distributed visual attention, pedestrians often did not even notice the eHMI on the car. Similarly, in a Wizard of Oz study in real traffic, Shuto et al. (2018) reported that pedestrians did not glance at the AV more often when eHMI was present as compared to when it was absent.

Apart from visual attention requirements, eHMIs may cause confusion because it may be unclear which road user the eHMI message addresses (for discussion, see Tabone et al., 2021a) or because the eHMI message is not intuitive (see Ackermann et al., 2019; De Clercq et al., 2019; Hensch et al., 2019). The literature is replete with examples of confusing eHMI messages, such as the text ‘GO’ (Eisma et al., 2021), arrows (Kunst et al., 2022; Zang et al., in press), and a red lamp (Bazilinskyy et al., 2020) being misunderstood regarding whether the message represents an instruction for the pedestrian or a representation of the AV’s intent. Also, some light-based eHMIs have been misinterpreted as sensors instead of communication devices (Bazilinskyy et al., 2019; Fratini et al., 2021; Shuto et al., 2018).

The good news is that research suggests that pedestrians get easily accustomed to novel types of eHMIs after a number of encounters (e.g., Colley et al., 2022; De Clercq et al., 2019; Eisele and Petzoldt, 2022; Hochman et al., 2020; Lee et al., 2022). However, a question remains whether, in conditions of time pressure or stress, participants would refer back to what is intuitive rather than learned (cf. Taylor and Garvey, 1959) or fail to see the eHMI altogether (see Bazilinskyy et al., 2022 for the use of eHMIs in near-collision situations). Furthermore, a pitfall is that repeated exposure to eHMIs can cause overreliance, resulting in dangerous interactions. An example of this is offered by Kaleefathullah et al. (in press). These authors let pedestrians cross a road in an immersive virtual environment. The pedestrians encountered AVs with a LED-strip eHMI that signaled that the AV would stop. At the 19th encounter, the LED strip turned on, but the AV did not stop. The results showed that many pedestrians stepped onto the road or even walked under the virtual car (see Fig. 1). Kaleefathullah et al. argued that such apparent eHMI failures might also occur in reality, for example, when the AV fails to detect the pedestrian but stops for another pedestrian further down the road.

3. Arguments for eHMIs

3.1. Argument 1 for eHMIs: eHMIs can contribute to ‘superhuman performance’

The first argument in favor of eHMIs is referred to as ‘superhuman performance’. To explain this, it is useful to first remember how an AV works: An AV senses elements of the road environment via cameras, radar, etc., and classifies objects using computer vision. The next step in the control loop is that the robot analyses the sensed data and makes plans and decisions. Next, the robot implements actions; that is, it controls itself and acts upon the environment. These functions are equivalent to the stages of automation as outlined by Parasuraman et al. (2000).

The essence here is that the robot (AVs) has knowledge about the environment and has plans about what it will do in that environment. For example, the AV likely knows where other road users are, what route to drive, and when to slow down for upcoming curves or crosswalks. In fact, path planning is a key subfield of robotics (Latombe, 2012; Sucan et al., 2012). Given that an AV has knowledge about the state of the environment and its future actions, the AV could share its upcoming actions using an eHMI. Such sharing would open up possibilities that may increase the efficiency and safety of the traffic. That is, the traffic performance could be higher than it is now, something that we call superhuman performance.

By comparison, in today’s non-automated traffic, drivers also have knowledge about the environment and plans about which maneuvers they will perform. However, this knowledge is fuzzy and private to the driver. While drivers can indicate direction, turn their head and eyes, use the horn, or flash high beams, they have limited capacity to indicate, for example, whether they have the intention to maintain or reduce speed. Today’s means of explicit communication, such as turn signals, brake lights, high beams, and the horn, are easily operated by a human driver (by pressing the brake, pulling/turning a lever, or pushing the steering wheel) but are not necessarily optimal for future automated and connected traffic.

De Clercq et al. (2019) tested a number of eHMIs on the front of the AV (e.g., front brake light, smiling display, text ‘WALK’, moving LED strip) in a virtual environment presented from a pedestrian’s perspective. The experimenters asked participants to hold a button whenever they felt safe to cross for approaching AVs that stopped or maintained speed. The results indicated that each of the eHMIs improved performance compared to no eHMI. That is, the eHMIs made participants feel safe to cross the road when it was indeed safe to cross (i.e., when the car stopped) and less safe to cross the road when it was unsafe to cross (i.e.,
when the AV maintained speed).

Fig. 2 provides the corresponding results for AVs that came to a full stop. It shows the percentage of participants holding the button as a function of the distance between the pedestrian and the AV. In all cases, the car started braking when it was 35 m from the pedestrian. The eHMI switched state from non-yielding to yielding at 50 m (before the car started braking), at 35 m (the same moment the car started braking), or at 20 m (after the car started braking). The black dashed line represents
the button-press results for the baseline condition without eHMI. It can clearly be seen that after the eHMI had switched state, participants felt safer to cross, i.e., the blue surface area lies above the black dashed line. Importantly, the biggest gains were achieved for the 50-m condition, that is, when the eHMI switched on before the vehicle braked.

The results in Fig. 2 illustrate that eHMIs are not to be seen as a replacement for something that is missing. Rather, eHMIs should be seen as an augmentation of current traffic and as a means of ‘making visible the invisible’. In this way, a level of performance is attained that is higher than when relying on implicit communication only.

3.2. Argument 2 for eHMIs: Pedestrians want eHMIs

The previous argument addressed performance, i.e., does an eHMI make pedestrians cross when they can and does it inhibit pedestrians from crossing when it is unsafe to cross? The current argument is about acceptance. does pedestrians want to receive information from an eHMI? Acceptance can be measured via questionnaires alongside an experiment. For example, in the above-reviewed experiment of De Clercq et al. (2019), participants were asked to rank the five experimental conditions (four eHMI types and a baseline condition). The baseline condition came out worst, with 23 of 28 participants indicating this as the least preferred choice among the five options.

Many other virtual-reality studies confirm high acceptance ratings of eHMIs (e.g., Deb et al., 2018; Dou et al., 2021; Ferenczak and Shafique, 2022; He et al., 2021). These findings are supported by a number of evaluations in real traffic:

- In the CityMobil2 project, pedestrians encountered AVs during demonstrations in three European cities. Questionnaires revealed that pedestrians would appreciate receiving information from the AV, for example, about whether it will stop, its speed, and whether it has detected the pedestrian (Merat et al., 2018).
- Cefkin et al. (2019) tested an eHMI in the form of a light strip indicating the intention of the AV in a busy urban environment. In post-experiment interviews, participants reported that they did not notice the eHMI. However, most participants also reported that they liked the idea of an eHMI and that they would pay more attention if they were more used to it.
- Monzel et al. (2021) evaluated an eHMI in the form of a front brake light mounted on 102 vehicles driving at an airport for 3.5 months. Questionnaires and interviews with 197 staff members who encountered the vehicles revealed that the eHMI was moderately positively received (with a mean frequency of positive and negative experiences of 3.35 and 1.99, respectively, on a scale of 1 = never to 5 = very often), where it was regarded as a means to better understand and predict the behavior of the vehicles.
- Forke et al. (2021) conducted a field study in mixed traffic, where pedestrians encountered an automated shuttle indicating its intention and awareness of other road users by means of an eHMI. In post-experiment interviews, 17 participants indicated that the external communication was necessary, 8 indicated it was helpful but not necessary, and 5 indicated it was not necessary. Furthermore, 16 participants indicated that the eHMI increased comprehensibility or predictability, whereas 14 reported no difference.

3.3. Argument 3 for eHMIs: Different forms of eHMI signaling are possible

The third argument in favor of eHMIs is that eHMIs do not just have to provide an instruction or show their intention, as was the case in the above-reviewed experiment of De Clercq et al. (2019) and the majority of eHMI research so far (Dey et al., 2020a). In a review paper by Schieben et al. (2019), four strategies for eHMI communication are outlined:

1. Information about the AV’s driving mode (manual/automated).
2. Information about the AV’s maneuvers or upcoming maneuvers (as already explained).
3. Information about AV’s perception of the environment.
4. Information about AV’s cooperation abilities.

Regarding the first communication strategy, it is known that mode confusions are a common issue in human-automation interaction. In aviation, many accidents have been attributed to mode errors (Mumaw, 2021; Sarter and Woods, 1995; Silva and Hansman, 2015), where the pilot executed an action appropriate for one mode while the automation system was actually in a different mode. Mode errors have been attributed not only to the large number of modes but also to the interaction between modes, i.e., mode changes can be initiated by the pilot, as well as indirectly through environmental triggers or completed tasks (e.g., target altitude achieved). In aviation, the human-machine interface plays a vital role in preventing mode confusion (Sarter and Woods, 1995).

In the same vein, in automated driving currently on the road, one sees a variety of automation subsystems, such as adaptive cruise control, automated lane keeping, automated emergency braking, and traffic light detection. These systems combined constitute the ‘automated driving system’. It is easy to imagine that drivers may have difficulty understanding which automation system is active (Banks et al., 2018; Dönmez Ozkan et al., 2021; Feldbittner et al., 2017). Similarly, other road users, such as pedestrians, may have difficulty understanding whether an approaching vehicle is driving automatically or not and whether this vehicle can be expected to respond to their presence. An eHMI could resolve this confusion by communicating the current automation mode.

Information about the AV mode is especially important in situations where clarity is needed about who is in control and about what can be expected from the car occupants. For example, confusion may arise if the car behaves indecisively (e.g., in a deadlock situation) or if the driver in the car is not paying attention (e.g., making a phone call). In the latter case, a police officer will not have to issue a fine if it is known that the car is driving automatically at that time.

Regarding the third point, AVs at present have imperfect sensing abilities. It is hard for a computer vision system to anticipate whether a pedestrian standing on the curb is about to cross the road or not, something that should be inferred from the pedestrian’s posture or bodily signals (Rudenko et al., 2020). Whether the AV understands the traffic situation could be conveyed via an eHMI. For example, Epke et al. (2021) tested a system where the driver gestured towards a simulated AV, and the AV signaled back I SEE YOU (Fig. 3). Epke et al. showed that the confirmatory message allowed the pedestrian to terminate the gesture earlier in time, as the pedestrian knew that the AV understood the gesture. Similarly, Colley et al. (2021) showed that confirmatory eHMI messages (‘Thank you’ or ‘You’re welcome’ as a response to, respectively, a pedestrian’s gesture giving right of way to the AV or thanking the AV for stopping) increased pedestrians’ trust in the AV as well as perceived safety and intelligence as compared to no confirmatory messages.

Regarding the fourth point, although research on HMI for cooperative driving exists (e.g., Zimmermann et al., 2015), research into eHMIs for cooperation in traffic is still scarce. The ISO guidelines provide one example in which road users cooperate: “an ADS-DV [Automated Driving System – Dedicated Vehicle] that is approaching the crosswalk can display to the car behind that there are pedestrians crossing the road (something that the vehicle behind cannot ‘see’).” A similar concept was proposed by Ter Borg et al. (2019), where a parked car warns a pedestrian about an approaching vehicle (Fig. 4). In theory, eHMIs could facilitate any type of multi-agent interaction, using high-definition maps and V2X communication as inputs.

In summary, eHMIs could communicate not only the AV’s maneuver intentions but also the automation mode and whether the AV understands the situation.
3.4. Argument 4 for eHMIs: Eye contact is important

Moore et al. (2019) suggested that eHMIs are unneeded because eye contact does not play a major role in traffic. Although eye contact is not essential, it could still be a feature that helps predict how a situation will unfold. In a recent experiment in a parking garage with 36 participants (De Winter et al., 2021), it was measured using a head-mounted eye-tracker where pedestrians look when walking around. The results showed that pedestrians often glanced at the driver (see Fig. 5, for an example), presumably to predict what the driver would do, infer whether the driver had seen them, or communicate that they have seen the driver. Pedestrians did not look at the driver only; many other features were glanced at, such as the wheels and backs of cars. The eye-tracking analysis suggested that pedestrians appear to extract whatever information they may need to predict what will happen next.

Similarly, Uttley et al. (2020) investigated road user interactions in a car park and found that pedestrians looked at the drivers in 65% of the vehicle–pedestrian encounters as identified by two observers. Moreover, it was found that in the encounters in which the pedestrian did not look at the driver, it was more likely that the driver slowed down without coming to a complete stop, whereas when the pedestrian did look at the driver, the driver either stopped or continued. The authors argued that the slowing down was indicative of the driver’s hesitation and that looking at the driver “is prompting the driver to make a clear decision about their behaviour” (p. 42).

The two studies discussed above were conducted in a parking garage/car park, which is a relatively unstructured environment. However, a recent online study suggests that eye contact is also relevant in a more straightforward scenario of a single vehicle approaching a single pedestrian (Onkhar et al., 2022). In Onkhar et al., the onset and offset timing of the driver’s eye contact and yielding behavior of the vehicle were varied in different configurations. Participants had to hold a key on the keyboard whenever they believed it was safe to cross and answer post-trial questions on the intuitiveness of the driver’s eye contact. The conclusions of the study by Onkhar et al. (2022) were as follows:

(1) Implicit communication is dominant. In simple words, pedestrians do not want to cross in front of a car that drives off, even if the driver in the car makes eye contact.

(2) Driver eye contact has no effect when the car is still far away, and eye contact cannot be seen.

(3) Pedestrians can still figure out what to do when the driver makes no eye contact at all.

(4) However, driver eye contact has a strong effect in a specific zone. That is, when the car is slowing down or waiting, eye-contact increases pedestrians’ willingness to cross and ensures a more intuitive interaction. Qualitatively, this zone corresponds to what Dey et al. (2021a) referred to as “when the speed of the vehicle is slow enough to not be an obvious threat, but still fast enough to raise a doubt about a vehicle’s stopping intention”.

In summary, eye contact does appear to have a role in traffic. Eye contact is not the only cue that pedestrians use; the importance of eye contact is likely dependent on many factors, including the distance to the approaching vehicle (see also Dey et al., 2019) and whether other
cues (e.g., traffic lights, implicit communication) provide redundant information. Consequently, when an AV features a distracted driver or no driver at all, this is not disastrous, but it may create some uncertainty or confusion on behalf of the pedestrian. The implication is that eHMIs may be needed to fill that social interaction void.

### 3.5. Argument 5 for eHMIs. Implicit communication causes confusion

The fifth and final argument for eHMIs is about implicit communication. First, consider Fig. 6, which introduces four different forms of implicit and explicit communication:

- **Left top:** Implicit communication can be non-anthropomorphic, i.e., *not* as a human would communicate. For example, an AV could change its pitch (i.e., forward tilt) or move laterally in the lane in a way that is not human-like in order to communicate its yielding intention (for studies into the effects of different forms of lateral movement, see Fuest et al., 2018; Sripada et al., 2021, and for pitch, see Bindschadler et al., 2022; Dietrich et al., 2019; Othersen et al., 2019b). Schmidt et al. (2019) evaluated a wide range of ‘engineered vehicle trajectories’ for AV-pedestrian communication. The trajectories differed in terms of approach speed, acceleration, deceleration, and pedestrian responsiveness, and many were non-anthropomorphic, i.e., not according to common social conventions.

- **Right top:** Non-anthropomorphic communication can also be explicit, i.e., in the form of eHMIs, as discussed in this work.

- **Left bottom:** It is also possible to conceive anthropomorphic implicit communication, for example, an AV that crawls forward as a human would to indicate that it wants to go first (Bazilinsky et al., 2021; Niedermeyer, 2019), an AV that stops before the stop line like a human would (cf. Domeyer et al., 2019; Risto et al., 2017; Schmitt et al., 2022), or an AV that keeps a human-like safety margin from objects and road users (Kolekar et al., 2021). In fact, an AV could be designed so that its driving behavior is indistinguishable from manual driving, where the AV passes the Turing test of automated driving (Emuna et al., 2020; Li et al., 2018; Stanton et al., 2020).

- **Right bottom:** Finally, one may consider explicit anthropomorphic communication, such as eHMIs in the form of a smile/smiling face (Deb et al., 2018; Dou et al., 2021; Joistien et al., 2021; Lücken et al., 2019; Praticci et al., 2021), artificial eyes (Chang et al., 2019; Lücken et al., 2019; Morales et al. 2020; Verstegen et al., 2021), an animated face (Bai et al., 2021; Mahadevan et al., 2018), a robotic hand (Mahadevan et al., 2018; Zhang et al., in press), visual embodiments of a driver (Furuya et al., 2021), or a humanoid steering robot (Mirnig et al., 2017).

Although implicit communication may be a dominant cue in traffic, it too can be a source of confusion and error. The literature contains many examples where pedestrians over- or underestimate vehicle speed and time gaps (Lobois and Cavallo, 2007; Papić et al., 2020; Sun et al., 2015). In fact, it can be argued that in current traffic, accidents often happen because vehicle communication is unclear, i.e., a failure to anticipate what the car will do. In a study analyzing accidents across Europe, Habibovic and Davidson (2012) found that in 36 out of the 56 crashes with vulnerable road users (20 pedestrians and 36 cyclists) at intersections, the vulnerable road user did see the vehicle but misjudged the situation. For example, they erroneously thought that the driver had seen them and would adjust to their presence, or they misjudged the timing of their crossing (for similar findings on cyclists, see Räsanen and Summala, 1998). A possible solution to these incidences in the case of automated driving would be to develop AVs capable of detecting vulnerable road users reliably (so that the AV slows down) or to ensure that the AV indicates its turning intentions early via an eHMI.

AV researchers are now attempting to create AVs that drive like humans (e.g., Fu et al., 2019; Wang et al., 2020; Zhu et al., 2018) (Fig. 6, left bottom). The assumption here is that such communication is most clear and transparent to the AV occupants and outside VRUs (Hecker et al., 2019). However, whether AVs should attempt to drive like humans is still an open question. In this context, it is useful to highlight the findings of a study in which pedestrians rated an AV that drove past (Bazilinsky et al., 2021). The AV drove in different ways: ‘playback manual’, but also ‘stereotype automated driving’ in the form of driving closer to the lane center, and ‘stereotype manual driving’ in the form of an AV that cut curves. The results showed that driving more towards the center, i.e., ‘stereotype automated driving’, yielded likeability ratings of pedestrians on par or slightly higher than playback manual driving. From this study, it was concluded that AVs do not have to drive like a human in order to be liked. It was also found that particular vehicle behaviors, such as crawling forward, can be prone to misinterpretation. Pedestrians may attribute such behavior to human failure (did the driver make an error?) or computer intelligence (is the AV programmed to drive like a human?).

The previous paragraph was about anthropomorphic implicit communication. Research has also been performed on non-anthropomorphic implicit communication (Fig. 6, left top). Recently, Sripada et al. (2021) presented participants with animated video clips of a car that showed a lateral deviation in the lane to indicate stopping intention, a study done with 1104 participants. Different left–right mappings and degrees of lateral deviation were investigated. Overall, it was found that when the car moved towards the participant (pedestrian), this made participants think that the car would stop for them. However, the effects compared to no lateral deviation were rather small. A noteworthy finding was that only a small portion of the participants appeared to understand the lateral motion. Many participants thought that the approaching vehicle reacted to their presence, i.e., pedestrian avoidance, or believed that the lateral deviation occurred because of an error by either a human driver or a hardware failure of the AV.

The bottom line is that implicit communication is not a panacea. Implicit communication can be misperceived and misinterpreted, and attempts to let AVs drive like a human or perform non-human-like maneuvers may confuse pedestrians. These observations suggest that implicit communication by AVs may require standardization and training to be understood, just like eHMIs need to.

### 4. Discussion

In this work, we presented common lines of reasoning for and against eHMIs. In short, arguments against eHMIs are that (1) there is no social interaction void to be filled because vehicle movement is dominant, (2) there is a wide variety of eHMI concepts and unresolvable dilemmas about how to meaningfully proceed, (3) eHMIs can have various negative effects as they have to compete for pedestrians’ limited visual attention. Negative effects also include confusion and overreliance.

Arguments for eHMIs are that (1) eHMIs can complement implicit communication, resulting in ‘superhuman performance’, (2) eHMIs are something that road users seem to want and accept, (3) eHMIs are not limited to communication of stop-and-go intentions/instructions; an eHMI can also indicate whether the AV’s sensors are functioning correctly and whether the automated driving systems are currently active, (4) eye contact does play an important role in current traffic, which suggests that eHMIs need to fill the social interaction void that is...
arises in driverless vehicles, (5) implicit communication can be misperceived and cause confusion, suggesting that pedestrians may benefit from explicit communication by eHMs.

This work aimed to advance the discussion on the need for eHMs by forwarding several common and novel arguments. However, this work does not provide the final answer regarding whether eHMs should or will be used in future traffic. Although this paper discussed a multitude of arguments, some factors were not considered. One of them concerns the effect of the recognizability of the AV, i.e., how external sensors (e.g., lidar on the roof), ‘self-driving’ stickers, and the presence/visibility of a driver affect pedestrian-AV interaction, and how these vehicle features interact with the eHMI (e.g., Vlakveld et al., 2020). Furthermore, the issue of liability was not discussed. It can be imagined that manufacturers would like to install eHMs on their AVs to ensure a transparent interaction. Another factor is that the field of automated driving continues to evolve. The capability of the AV, such as whether the AV travels on a segregated road or can enter more complex environments, is undoubtedly important for the type of eHMI communication required.

In that sense, the field of Human Factors and eHMs can be expected to evolve along with technology. It is hoped that the present paper will help advance the discussion on the need for eHMs in future traffic.

CRediT authorship contribution statement

Joost de Winter: Conceptualization, Writing – original draft, Funding acquisition. Dimitra Dodou: Resources, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research is supported by grant 016.Vidi.179.047 (“How should automated vehicles communicate with other road users?”), which is financed by the Netherlands Organisation for Scientific Research (NWO).

References


Ahn, S., Lim, D., Kim, B., 2021. Comparative study on differences in user reaction by forwarding several common and novel arguments. However, this work that sense, the field of Human Factors and eHMIs can be expected to evolve along with technology. It is hoped that the present paper will help advance the discussion on the need for eHMs in future traffic.


