Interpreting externally expressed intentions of an autonomous vehicle

Maarika Oidekivi
Institute of Technology
University of Tartu
Tartu, Estonia
maarika.oidekivi@ut.ee

Alexander Nolte
Institute of Computer Science
University of Tartu
Tartu, Estonia
alexander.nolte@ut.ee

Alvo Aabloo
Institute of Technology
University of Tartu
Tartu, Estonia
alvo.aabloo@ut.ee

Karl Kruusamäe
Institute of Technology
University of Tartu
Tartu, Estonia
karl.kruusamae@ut.ee

***Abstract*—With the imminent addition of autonomous vehicles to traffic, it is becoming more vital to look at different alternatives to non-verbal communication between the driver and the pedestrian, so that the pedestrian would understand the intentions of autonomous vehicles. The aim of this paper is to evaluate different approaches to communicate the intent of an autonomous vehicle. A survey study was conducted among Estonian people to analyze their understanding of animations of more prominent explicit external interaction modalities. The study revealed that participants may not understand the vehicle intent if they have no prior knowledge about the displayed signals.**

***Index Terms*—human-vehicle interaction (HVI), autonomous vehicle, machine intent, explicit HVI, external HVI, social robotics, user-study**

I. Introduction

As autonomous technologies are integrated into everyday traffic, the question arises whether current road users (e.g. drivers of regular vehicles, pedestrians, and passengers of robotaxis) are able to interact with these autonomous agents. There is a strong case that in addition to accurately interpreting its environment, the autonomous vehicle should externally convey its intent (e.g. near-term future actions) to other road users in order to ensure smooth traffic flow and safety. However, there is no consensus in the community of human-vehicle interaction whether this intent should always be conveyed explicitly [1], [2]. Road use is already quite stringently regulated by vehicle- and infrastructure-based explicit signaling (e.g., turn signals and traffic light). Thus the ongoing debate is mostly about whether recognition and expression of informal communicative cues such as hand gestures and eye contact should also be integrated for the autonomous road users [2].

Most often the case for external human-vehicle interaction is argued from the perspective of a pedestrian crossing the road when an autonomous car is approaching (e.g., [1]–[4]). Arguably it becomes relevant for the pedestrian to understand the intentions of the autonomous vehicle before deciding if it is safe to step into the road. Approximately 90% of the pedestrians look at the approaching vehicle before crossing the road [5]. Moreover, most pedestrians and drivers engage in some form of implicit interaction and there are state regulations that even require this behavior. For instance, Estonian Traffic Act stipulates that for safe crossing, "a pedestrian must take into account the distance and speed of approaching vehicles before stepping onto the carriageway, give the drivers an opportunity to smoothly reduce the speed or bring the vehicle to a halt, and make sure that the driver has noticed the pedestrian" [6].

The expression of external intent by a vehicle can be either explicit or implicit [1]. In case of explicit interaction, the vehicle uses clearly visible displays, light panels, and/or projections whereas implicit expression makes use of vehicle's change of speed or engine sounds to convey intent [1].

Explicit cues provide a clear and potentially unambiguous way of expressing the intent of a vehicle. The majority of empirical studies show higher efficiency and perceived safety when a vehicle is equipped with an explicit communication interface [2]. Nevertheless, several drawbacks have also been identified [1]. For example, light patterns and projections are difficult to see in bright sun and screens can distract pedestrians from other important signs of danger or confuse them [1].

The interface used to express intent should be designed to be simple, understandable and quickly accepted by pedestrians. There is a notable amount of prototype development by car manufacturers and autonomy providers to add different types of displays to convey the machine's intent but controlled evaluations of proposed systems are rare until now [2].

The research focus of this paper is thus to investigate how well people interpret the most common explicit signals proposed by industry leaders. For this purpose we designed an online survey where participants are exposed to animations of the more prominent explicit external interaction modalities by an autonomous vehicle during pedestrian crossing. The main objective is to evaluate how likely are the proposed modalities correctly interpreted by an uninitiated pedestrian. And as analogous surveys are needed to map any culturally-driven similarities and differences, the study adds results from the Estonian-speaking setting to an emergent field of human-vehicle interaction.

This study was in part supported by European Social Fund via Smart Specialization project with Bolt Technologies OÜ and Estonian Centre of Excellence in IT (EXCITE) funded by the European Regional Development Fund.

978-1-6654-4112-4/21/$31.00 ©2021 IEEE

II. Related work

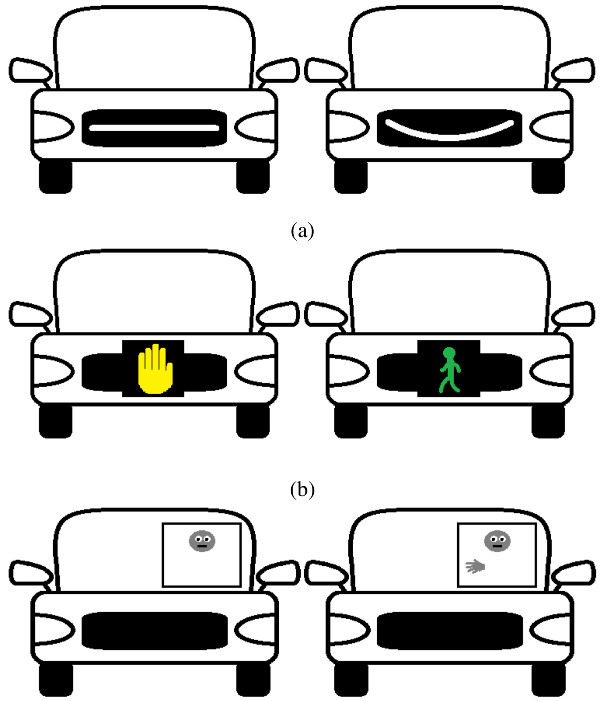
In this section we will review some notable examples (LED displays, LED strips, and anthropomorphic features) along with key findings from respective user studies. This section draws relevant examples from the global context to position our survey design and its eventual results. While the section highlights the characteristic technological approaches and the associated empirical results, a more comprehensive overview can be found in a review article by Rouchitsas & Alm [2].

A. LED displays

Matthews et al. proposed a system that used an LED display to explicitly give a written instruction to the pedestrian (Fig. 1) [7]. It was considered important for the people to feel comfortable near the new technology, therefore human-like features were avoided and only familiar features were used. The displayed texts were kept clear and short to avoid confusion. The experiments showed that the intent communication system shortened interaction times between the pedestrians and the vehicle and it turned out that the level of human comfort around the vehicle depended on how much prior information they had about the vehicle and how far they were from the vehicle. [7]

Fig. 1: System used by Matthews et al. Reprinted from [7] under CC BY-NC-SA 4.0.

Holländer et al. studied three different LED display concepts, all using two states: neutral and yielding [8]. Firstly Semco's smiling car was used (Fig. 2a). A horizontal bar was shown when the vehicle was in motion and did not intend to stop. If the system intended to let pedestrians cross, the horizontal bar was animated to a smile. Secondly the "green man"/"yellow hand" concept (originally proposed by Friendman et al. [9]) was used. A green man was displayed when pedestrians could cross the road and a yellow hand was displayed when they had to wait (Fig. 2b). Both the smiling car and the "green man"/"yellow hand" were located at the grille. As the third concept, an animation of a robot was placed on the driver's side of the windshield (Fig. 2c). The animated robot only interacted with pedestrians if the vehicle intended to stop: it waved its hand from left to right indicating it was safe to cross. If the vehicle continued to drive, a static image of the robot's face was presented. The study showed that external car displays influenced pedestrians' decision-making. The participants reacted faster with an external display compared to an inactive external display. From the presented methods, the "green man/yellow hand" representation performed the best.



(a)

(b)

(c)

Fig. 2: Concepts investigated by Holländer et al: (a) smiling car, (b) colored pictographs, (c) animated robot [8].

Razmi Rad et al. created a concept, where vehicle intent was shown by altering the color of the hood of a vehicle (Fig. 3) [10]. In the first case the interface turned yellow when a pedestrian was noticed and the vehicle intended to yield. In the second case yielding was indicated by green and not yielding by red. In addition, the relationship between pedestrian behavioral habit and their behavior near autonomous vehicles was studied. It turned out that the combination of green and red color was generally better understood, 55% of the pedestrians made a correct road crossing decision in that case. The behavior study found that pedestrians who usually follow the rules were more likely to stop and let the autonomous vehicle drive, the pedestrians with more risky behavior in traffic may start taking advantage of the autonomous vehicle's detection system and cross the road with less attention [10].

B. LED light strips

Habibovic et al. designed a LED light strip (Fig. 4) with four modes:

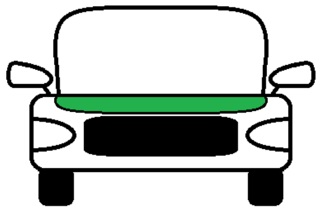


Fig. 3: Concept by Razmi Rad et al. [10].

- the vehicle is in automated mode,
- the vehicle is about to yield,
- the vehicle is waiting,
- the vehicle is about to start driving [11].

The autonomous vehicle's intent was shown by changing the frequency and the area of the signal. To avoid confusion with signals already used in traffic, red, green, blue and amber were excluded and white/yellow was used for communicating all messages. The experiments revealed that without having previous knowledge only a few participants could identify the autonomous vehicle's intent. After some explanation, the participants could understand the signals and found them to be easy to understand. [11]

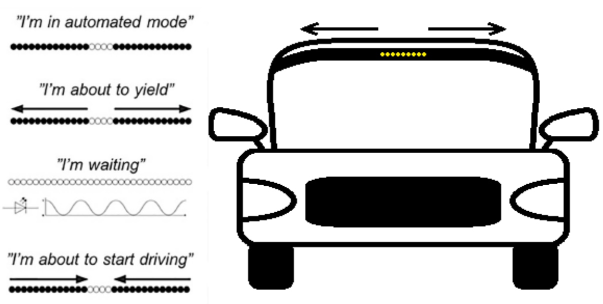


Fig. 4: Concept by Habibovic et al [11].

Benderius et al. used a "show, don't tell" principle in their LED strip [12]. This means that the vehicle will always clearly demonstrate its planned actions, but without any further motivation. For example, if a pedestrian and a cyclist will approach a zebra crossing at the same time, the vehicle will indicate that it is going to stop and wait, but it will not state which agent triggered this action. The LED strip has the following functions:

- indicating the distance from an object (shown by the width of the lit area),
- warning any conflicting lead vehicle (flashing light bar),
- indicating the intended direction of movement (position of the lit area). [12]

C. Eyes

Chang et al. used a car with mock eyes to generate eye contact between the car and a pedestrian (Fig. 5) [13]. The eyes looked at the pedestrian, when the vehicle intended to yield. A survey was used to establish that the area of headlights is the most suited for positioning the eyes. The experiments in VR concluded that 66,6% of the participants made the right road crossing decision quicker in case of a vehicle with the eyes even without any previous explanation. This percentage increased to 86,6% when the participants were asked to take notice of the eyes. The interview data showed that the eyes helped the pedestrians to recognize the vehicle's awareness of them and that the eyes interface could help reassure the pedestrian's street-crossing decision. [13]

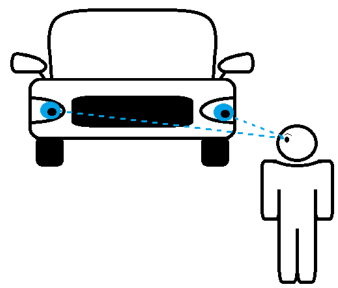


Fig. 5: Chang et al's car with eyes [13].

III. Methodology

To gain knowledge about the perception of autonomous vehicle's intent we opted to conduct an online survey. This approach is reasonable because our aim was to assess how well people would interpret the most common explicit signals proposed by the industry leaders on a broader scale.

The study was approved by the Research Ethics Committee of the University of Tartu (approval 299/T-27).

The survey consisted of two parts:

- 8 questions about interpreting vehicle's intent,
- 5 questions capturing the demographics (age and gender) as well as the technological background of the participants.

The participants are asked to imagine that they are about to cross the road and there is a car that is either (1) approaching or (2) already at the pedestrian crossing. For each scenario, four questions were developed. Every question is posed as a text accompanied by a photorealistic animated GIF depicting the expression of the intent. The participant is requested to select the response from a list of 5 answers. The list of answers is identical within all the questions of the same scenario.

For the first scenario the question and the list of available answers are as follows:

"Imagine that you are about to cross the road and there is a car approaching that displays the following. How would you interpret the situation?"

1) It is a decoration with no significance to a pedestrian.
2) The display indicates that the vehicle is stopping to yield.

3) The display indicates that the vehicle is planning to continue driving, there will be no change in its behaviour.
4) I guess the displayed information might be intended for a pedestrian but I do not understand its meaning.
5) Other.

For the second scenario the question and the list of available answers are as follows:

"Imagine that you are about to cross the road and there is a car stopped in front of the crossing that displays the following. How would you interpret the situation?"

1) It is a decoration with no significance to a pedestrian.
2) The display indicates that the vehicle is yielding.
3) The display indicates that the vehicle is about to start moving.
4) I guess the displayed information might be intended for a pedestrian but I do not understand its meaning.
5) Other.

The photo-realistic animated GIFs were generated based on two concepts proposed by Mercedes-Benz (Fig. 6) and Ford (Fig. 7). These two concepts were chosen for several reasons:

1) together they capture the most frequently proposed external human-vehicle interaction systems: LED lights strips and projections on the road [2];
2) Ford has proposed the LED light strip with this particular set of illumination patterns to be the industry standard [14];
3) there existed photorealistic animations of the proposed concepts which could be edited to display only a single intention and thus transformed into animated GIFs used in this study.

To collect the data an online survey tool, LimeSurvey, was used. The survey was distributed through mailing lists of different Estonian Universities. Descriptive statistics of the collected data were calculated using R (version 3.6.3) to analyze the distribution and aggregated accuracy of the responses.

The online survey, collected data, and the R scripts used to process the data can be found in the supplemental material available at: doi.org/10.5281/zenodo.4725710.

IV. Results

In total 171 people answered the survey, this included 113 women and 58 men between ages 18-30 years.

Fig. 8 shows how the participants interpreted the vehicle's intent when asked to imagine that they are about to cross the road and there is a vehicle approaching. The majority of participants (78%) correctly interpreted the vehicle's intent only in the case of Mercedes-Benz projecting on the road that it is about to yield (Fig. 8d). When LED light strips or panels were used to communicate either yielding or not yielding, the most common interpretation (around 50%) was that the lights are purely decorative and carry no significance to the pedestrian (Fig. 8a-c). The actual intent of the vehicle was recognized by less than 10% of the participants whereas roughly one quarter replied that they did not understand the meaning of the signal (Fig. 8a-c).

Fig. 6: Mercedes-Benz F 015 Luxury has LED panels in the front and back side of the vehicle. Additionally it is able to communicate with other road users by projecting visuals on the road [15].

Fig. 7: Ford has proposed a light bar on the windshield to communicate the intent of the vehicle [14].

Fig. 9 shows how the participants interpreted the vehicle's intent when asked to imagine that they are about to cross the road and there is a vehicle stopped by the zebra crossing. The results were similar to the first scenario. The majority of participants correctly interpreted the vehicle's intent in the case of Mercedes-Benz projecting on the road its intent to yield (76%) and in the case of Mercedes-Benz using the LED panel to express yielding (39%) (Fig. 9c-d). When LED strips were used to communicate yielding or that the vehicle is about to drive, people mostly thought that the lights were decorative and had no significance to the pedestrian (37 % and 48%) (Fig. 9a-b). In the case of LED strips, more than a quarter of the participants did not understand the meaning of the signal and the correct intent of the vehicle was recognized by 26% and 15% for about to drive and yielding, respectively (Fig. 9a-b).

V. Discussion

The survey of the autonomous vehicles' intent revealed that the participants did not understand the meaning conveyed by LED strips and displays. The participants mostly thought that

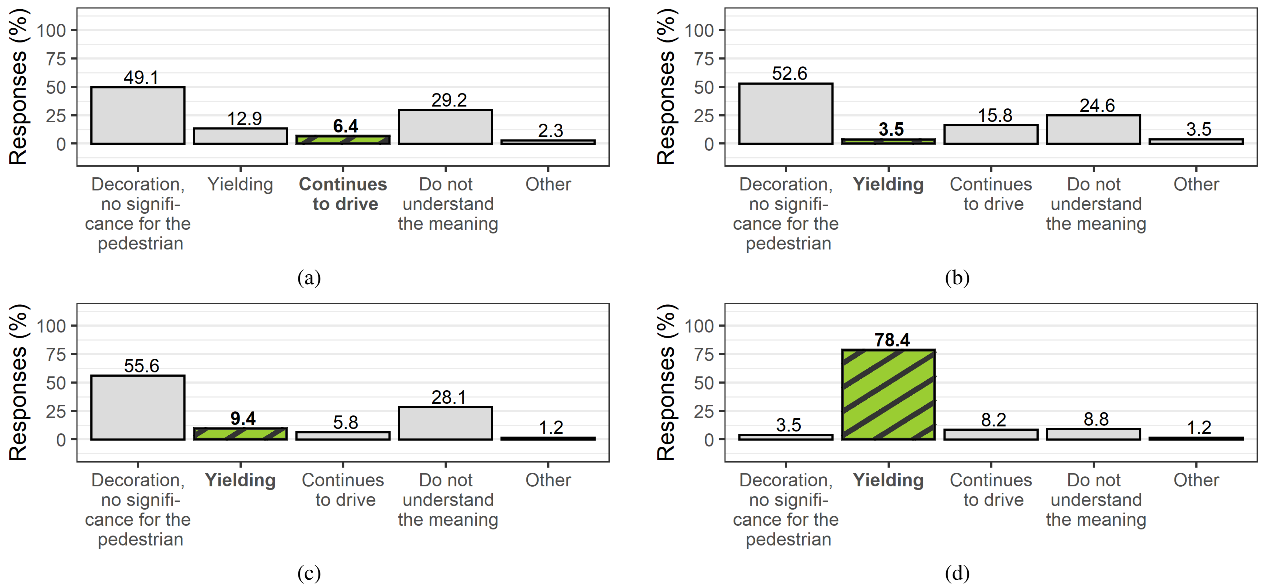


Fig. 8: Distribution of how the vehicle's intent was interpreted when the vehicle was approaching the pedestrian. The intention of the vehicle is highlighted in green. (a) Ford's LED light strip signaling it is continuing to drive; (b) Ford's LED light strip signaling it is stopping to yield; (c) Mercedes-Benz' LED panel signaling it is stopping to yield; (d) Mercedes-Benz projecting on the road it is stopping to yield.

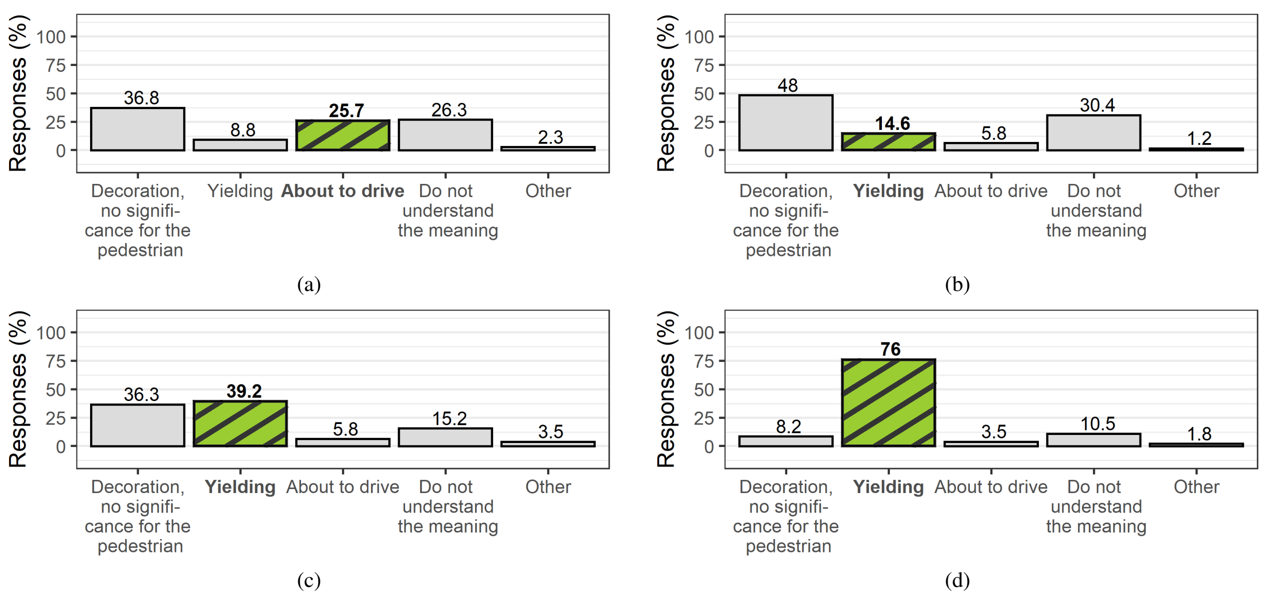


Fig. 9: Distribution of how the vehicle's intent was interpreted when the vehicle was stopping in front of the crossing. The intention of the vehicle is highlighted in green. (a) Ford's LED light strip signaling it is about to drive; (b) Ford's LED light strip signaling it is yielding; (c) Mercedes-Benz' LED panel signaling it is yielding; (d) Mercedes-Benz projecting on the road it is yielding.

the displays were meant as decorative elements that had no significance to the pedestrians or that the displays could be important to the pedestrians, but they did not understand their meaning. These results are in good agreement with the few previous studies [2], [16], [17]. Ackerman et al. [16] also concluded that automated vehicles should use the most familiar parameters, e.g., universal symbolism or simple text messages. Our results confirm that when new arbitrary signaling concepts are introduced, the pedestrian fail to interpret the signals correctly. Similar observation was made by Hensch et al. [17]. Nevertheless several studies indicate that even when signals are not correctly understood, majority of pedestrians prefer receiving explicit information about the vehicle's intent [2].

In our study, the vehicle intent was correctly interpreted only in the two cases where projections on the road were used to express the intent. This again corresponds to previous studies showing that without prior knowledge, people prefer projections and displays notably more than a LED light strip and that LED strips are perceived as not easily comprehensible [16]. It could be possibly due to the vehicle using already known elements, for example a zebra crossing.

The study design also poses some limitations. The participants were exposed to the intent signaling concepts without any prior instruction. In real world scenarios, the pedestrians are likely to encounter such signals on multiple occasions and inherently learn the meaning of even more abstract concepts. Additionally, as an online survey, there was minimal context provided for the participants. A potential Wizard-of-Oz type of study would bring a deeper sense of realism, which would increase the descriptiveness of the results.

In addition to signaling intent at a pedestrian crossing, future studies should also investigate the external human-vehicle interaction in other scenarios where pedestrians and drivers communicate. For instance, conveying intent when identifying and approaching passenger for a safe pick-up by a robotaxi.

VI. Conclusion

In this paper the results were presented from a survey study in which people were exposed to animations of the more prominent explicit external interaction modalities by an autonomous vehicle during pedestrian crossing.

The survey showed that people may not understand the meaning conveyed by LED strips and displays and for them to understand the displayed signals it may be necessary to have some prior knowledge.

References

[1] D. Moore, R. Currano, G. E. Strack, and D. Sirkin, "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 - AutomotiveUI '19*. Utrecht, Netherlands: ACM Press, 2019, pp. 295–307.

[2] A. Rouchitsas and H. Alm, "External human–machine interfaces for autonomous vehicle-to-pedestrian communication: A review of empirical work," *Frontiers in Psychology*, vol. 10, p. 2757, 2019.

[3] L. P. Robert, "The future of pedestrian-automated vehicle interactions," *XRDS*, vol. 25, no. 3, p. 30–33, Apr. 2019.

[4] A. Rasouli and J. K. Tsotsos, "Autonomous Vehicles That Interact With Pedestrians: A Survey of Theory and Practice," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 900–918, Mar. 2020.

[5] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Agreeing to cross: How drivers and pedestrians communicate," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. Los Angeles, CA, USA: IEEE, Jun. 2017, pp. 264–269.

[6] Road traffic act. Riigi Teataja. [Online]. Available: https://www.riigiteataja.ee/en/eli/ee/510072014013

[7] M. Matthews, G. Chowdhary, and E. Kieson, "Intent Communication between Autonomous Vehicles and Pedestrians," *arXiv:1708.07123 [cs]*, Aug. 2017, arXiv: 1708.07123. [Online]. Available: http://arxiv.org/abs/1708.07123

[8] K. Holländer, A. Colley, C. Mai, J. Häkkilä, F. Alt, and B. Pfleging, "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 - MobileHCI '19*. Taipei, Taiwan: ACM Press, 2019, pp. 1–11.

[9] L. Fridman, B. Mehler, L. Xia, Y. Yang, L. Y. Facusse, and B. Reimer, "To Walk or Not to Walk: Crowdsourced Assessment of External Vehicle-to-Pedestrian Displays," *arXiv:1707.02698 [cs]*, Jul. 2017, arXiv: 1707.02698. [Online]. Available: http://arxiv.org/abs/1707.02698

[10] S. Razmi Rad, G. Homem de Almeida Correia, and M. Hagenzieker, "Pedestrians' road crossing behaviour in front of automated vehicles: Results from a pedestrian simulation experiment using agent-based modelling," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 69, pp. 101–119, Feb. 2020.

[11] A. Habibovic, V. M. Lundgren, J. Andersson, M. Klingegård, T. Lagström, A. Sirkka, J. Fagerlönn, C. Edgren, R. Fredriksson, S. Krupenia, D. Saluäär, and P. Larsson, "Communicating Intent of Automated Vehicles to Pedestrians," *Frontiers in Psychology*, vol. 9, p. 1336, Aug. 2018.

[12] O. Benderius, C. Berger, and V. Malmsten Lundgren, "The Best Rated Human–Machine Interface Design for Autonomous Vehicles in the 2016 Grand Cooperative Driving Challenge," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 4, pp. 1302–1307, Apr. 2018.

[13] C.-M. Chang, K. Toda, D. Sakamoto, and T. Igarashi, "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 - AutomotiveUI '17*. Oldenburg, Germany: ACM Press, 2017, pp. 65–73.

[14] "Ford, Virginia Tech Go Undercover to Develop Signals That Enable Autonomous Vehicles to Communicate with People | Ford of Europe | Ford Media Center." [Online]. Available: https://media.ford.com/content/fordmedia/feu/en/news/2017/09/13/ford-virginia-tech-autonomous-vehicle-human-testing.html

[15] "The Mercedes-Benz F 015 Luxury in Motion." library Catalog: www.mercedes-benz.com. [Online]. Available: https://www.mercedes-benz.com/en/innovation/autonomous/research-vehicle-f-015-luxury-in-motion/

[16] C. Ackermann, M. Beggiato, S. Schubert, and J. F. Krems, "An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles?" *Applied ergonomics*, vol. 75, pp. 272–282, 2019.

[17] A.-C. Hensch, I. Neumann, M. Beggiato, J. Halama, and J. F. Krems, "How should automated vehicles communicate?–effects of a light-based communication approach in a wizard-of-oz study," in *International conference on applied human factors and ergonomics*. Springer, 2019, pp. 79–91.