arXiv:2307.12708v1 [cs.HC] 24 Jul 2023

Human-vehicle interaction for autonomous vehicles in crosswalk scenarios: Field experiments with pedestrians and passengers

R. Izquierdo¹, S. Martín¹, J. Alonso¹, I. Parra¹, M. A. Sotelo¹, D. Fernández-Llorca^{1,2}

This work has been accepted for publication at the 26^{nth} IEEE International Conference on Intelligent Transportation Systems.

Abstract— This paper presents the results of real-world testing of human-vehicle interactions with an autonomous vehicle equipped with internal and external Human Machine Interfaces (HMIs) in a crosswalk scenario. The internal and external HMIs were combined with implicit communication techniques using gentle and aggressive braking maneuvers in the crosswalk. Results have been collected in the form of questionnaires and measurable variables such as distance or speed when the pedestrian decides to cross. The questionnaires show that pedestrians feel safer when the external HMI or the gentle braking maneuver is used interchangeably, while the measured variables show that external HMI only helps in combination with the gentle braking maneuver. The questionnaires also show that internal HMI only improves passenger confidence in combination with the aggressive braking maneuver.

I. INTRODUCTION

Efficient human-vehicle interaction in the context of Autonomous Vehicles (AVs) has a fundamental impact on the user's sense of agency, perception of risk, and trust [1]. These factors, in turn, are essential to avoid both disuse and misuse of technology, which directly affect user acceptance and safety respectively [2].

Human-vehicle interaction in autonomous driving is a multi-user problem which, at a first level, includes mainly two groups of humans: those using the AV (i.e., passengers) and external road users interacting with the AV (i.e., pedestrians, cyclists, drivers). The lack of a driver to communicate with, both from the point of view of a passenger [3] and an external road agent [4], changes the nature and dynamics of interactions. In this new context, AVs need to communicate their intentions to non-automated or non-connected elements using all available resources. This communication is particularly important in scenarios where safety-relevant interactions occur, such as scenarios with pedestrians crossing in front of the AV.

The use of Vehicle-to-everything (V2X) [5] endows communications between other automated agents, such as other connected vehicles and the infrastructure, but still keeps humans unaware of the vehicle's intentions. Human-vehicle interaction is mainly developed through human-machine interfaces (HMIs), both internal (iHMI) and external (eHMI), whose specific modality is linked to vehicle technology and human capabilities [2]. The behavior of the vehicle, i.e. its movement dynamics, is also an important form of implicit



Fig. 1. Experiment configuration. The vehicle drives autonomously with a passenger towards the crosswalk while the pedestrian walks perpendicularly to the road and crosses.

communication with a considerable impact on the interaction [4], [6].

The impact of these forms of explicit or implicit communication on in-vehicle users (drivers or passengers) and other external road users, although widely studied, has always been done separately, which does not allow holistic conclusions to be drawn. From an experimental point of view, previous work focuses on simulated environments using virtual reality [7], or on real environments with two main types of constraints. On the one hand, cases in which the pedestrian only expresses an intention to cross without actually performing the crossing action [6]. On the other hand, cases where the driving is not really automatic but mediated by Wizard of Oz methods [8]. In all cases, the results are somehow limited due to the mismatch with real-world interaction scenarios.

In this work, we present the results of a real field study on human-vehicle interaction in crosswalk scenarios, involving both pedestrians and passengers. Our automated test vehicle [9] is not mediated by the Wizard of Oz approach. The pedestrians does not explicitly communicate their intention to cross (with the added difficulty of identifying the exact moment when the pedestrian decided to cross), but decide to cross (or not), and finalize their crossing action. In this way, on the one hand, we can draw conclusions that take into account the two types of users who interact with the AV in a holistic way. It also allows us to investigate the impact of previous experience of interaction as a passenger on pedestrian behavior and vice versa. On the other hand, we can minimize the gap between the interactions measured in our experimental setup and those that would occur in a real environment. We evaluate different types of internal and external HMIs, as well as implicit communication, and we use both behavioral and attitudinal methods.

¹Computer Engineering Department, Universidad de Alcalá, Alcalá de Henares, Spain. ruben.izquierdo@uah.es

²European Commission, Joint Research Centre, Seville, Spain.



(a) Screen showing no information. (b) Image shown at manual mode.



(c) Image shown at autonomous(d) Live video input with pedestrian mode. detection.

Fig. 2. Examples of the four possible statuses of the iHMI.

II. RELATED WORK

Studies used to assess human behavior in traffic scenes are mainly based on surveys which may not accurately collect the desired information [10]. Recording-based studies allow data to be obtained by direct measurement and allow certain biases to be eliminated [11]. Other technologies also allow recording eye-tracking [6] to analyze the focus of attention of the study subject and which factors are most relevant. In this study, both questionnaires and internal and external video recordings are used to analyze interactions.

In the eHMI section, we can find several features [12]: anthropomorphic, textual, light patterns or trajectory projection on the ground [13]–[15]. The color and shape of the communication play a fundamental role in the interaction. Messages using red color are more suitable to indicate risk, and green ones to indicate a safe situation [12]. This idea reinforces the theories of social constructivism [16]. However, in [6] a turquoise LED strip is used to convey different AV behaviors varying the light pattern but not the color. Turquoise color is commonly used as non-related-to-traffic color for "new" AV features.

The iHMI has the advantage of being integrated inside the vehicle cabin and can communicate information to passengers through visual and/or auditory clues. In the early stages, some prototypes were developed to share situational awareness with AV researchers [17]. However for highly automated vehicles and considering people inside the cabin as mere passengers new iHMI has been developed with the goal to improve user experience, acceptance, and trust [18]. The iHMI has been in deep study to facilitate transferring from autonomous to manual mode and vice versa. But, to the best of our knowledge, there are no studies that have evaluated the effect of an iHMI when simultaneously interacting with other road users in a safety-critical scenario. In this work, we present results collected from real experimentation with an AV driving autonomously and interacting with a pedestrian in a crosswalk using simultaneously an iHMI and an eHMI.



(a) GRAIL at solid green status.(b) GRAIL at solid red status.Fig. 3. Examples of the activated statuses of the eHMI.

III. EXPERIMENT DESCRIPTION

The experiment has the goal to create an interaction situation between an AV and its passenger with a pedestrian crossing through a crosswalk. The recreation of this situation under controlled conditions allows us to measure and evaluate their behavior through direct and indirect measurements and questions. This section provides a description of the recreated situation and different experiment configurations.

A. Experiment Configuration

Figure 1 shows a schematic representation of the experiment configuration. The pedestrian walks perpendicularly to the road with the goal to cross the road in the crosswalk. At the same time, the AV arrives at the crosswalk, and the interaction between the AV, its passenger, and the pedestrian takes place. The vehicle is driven autonomously along a straight stretch of road. The vehicle speed is 30 km/h until it starts the braking maneuver using a constant deceleration. The passenger is located in the co-drivers seat. In the driver's position, there is a backup driver to comply with current legislation but will not intervene unless critical and imminent risk. The pedestrian waits at the crosswalk with his back to the road until s/he receives the order to turn around and walk towards the pedestrian crossing. The position of the vehicle is used as a reference to trigger the pedestrian's reaction.

B. Test Variations

The autonomous vehicle has been equipped with two communication interfaces. The iHMI is a screen with the goal to communicate the current status of the vehicle to the passenger. Figure 2 shows the four possible states of the iHMI. Messages shown in figures from 2b to 2d are accompanied by audio messages describing the current status of the vehicle. The messages say: "Autonomous mode activated", "Autonomous mode deactivated", and "Pedestrian detected", respectively. The detection of the pedestrian is represented with a live-video stream accompanied by the detected bounding box of the pedestrian using a state-of-the-art CNN [19].

The eHMI, so-called GRAIL (Green Assistant Interfacing Light) [20] is a LED strip located in the front bumper and it is used to communicate the vehicle's intentions to other surrounding agents. Figures 3a and 3b show the two possible states in addition to shutdown. The GRAIL interface remains solid red when the vehicle drives at its cruise speed and turns solid green when it begins to brake and yields to pedestrians. In addition, we introduce other implicit forms of communication such as speed profiling. The way the vehicle decreases its speed and the anticipation of the maneuver have an important role to play when interacting with pedestrians in a crosswalk. With this goal in mind, we have defined two different braking maneuvers. One is denoted as gentle or optimal with low deceleration and high anticipation. On the other hand, the aggressive or reactive braking maneuver has a higher deceleration and lower anticipation. Both braking maneuvers start at 30 km/h and end with a full stop right before the crosswalk with two deceleration values, -0.9 m/s^2 for the gentle braking maneuver and -1.8 m/s^2 for the aggressive one.

C. Test Batch

Five tests have been defined to evaluate the perceived level of confidence and safety using different combinations of communication techniques. Table I shows the type of implicit and/or explicit ways of communication used for each test. Test number 0 is a drive-through the crosswalk area with no stop or speed reduction. The goal of this test is to prime both participants with the idea that there is a real risk and the vehicle may not stop. All tests were performed in a predefined random order except for test number 0, which was always performed in the first position.

 TABLE I

 CONFIGURATION OF EXPERIMENTATION TESTS

Test	Braking	Explicit		Stop
Number	Maneuver	Internal	External	
0	-	-	-	No
1	Gentle	-	-	Yes
2	Aggressive	-	-	Yes
3	Gentle	HMI	GRAIL	Yes
4	Aggressive	HMI	GRAIL	Yes

D. Experiment Sensing Setup

From the experiment point of view, different sources of information are needed to measure the interaction between the AV and the participants. The AV is equipped with an RTK-GPS system that provides precise positions of the vehicle with respect to the crosswalk. Two cameras complete the sensor setup used for the experimentation. The internal camera is mounted over the iHMI and records the passenger seat area. The external camera is mounted on the top of the vehicle and records the environment in front of the AV. The image of the pedestrian and its detection is shown in the iHMI when it is used.

E. Participants

Volunteers were recruited at the University area, including also family and friends. A total of 34 people joined the experiment but 2 of them could not complete the whole set and their information was discarded. The sample is N = 32 (18 men and 14 women) with an age distribution $\mu = 39.7$ and $\sigma = 12.6$ years.

Volunteers participated in the experiment as couples. One performed the passenger role while the other performed the pedestrian role. After finishing the complete set of tests, the participants swapped roles to perform the complete set of tests again in the same random order. This mechanism allows us to analyze whether the role initially played has any effect on the confidence perceived when interacting with the AV.

IV. EXPERIMENT EVALUATION

This section describes the tools used to assess participants' perceived confidence. For this purpose, a questionnaire was used with questions related to the perceived level of confidence, communication devices, and braking maneuvers. In addition, direct measurement variables were used during the experiments to evaluate not only what people think but also what they actually do.

A. Questionnaire

A questionnaire has been designed to ask participants about each interaction with the vehicle. As a pedestrian, three questions were answered after each experiment:

- Q1: What was your level of confidence that the vehicle would stop and yield to you?
- Q2: How did you perceive the braking of the vehicle?
- Q3: Has the visual communication interface improved your confidence to cross?

As a passenger, four questions were answered:

- Q4: What has been your confidence in the vehicle?
- Q5: How did you perceive the braking of the vehicle?
- Q6: Has the audiovisual communication interface improved the level of confidence in the vehicle?
- Q7: Which signal was most helpful to you?

Answers for questions from Q1 to Q6 are tabulated in a 7-step Likert scale [21] and three answers are possible for question Q7: "visual", "audio", or "both". The answers to these questions allow us to know how participants perceived the interaction with the vehicle as passengers and as pedestrians. Note that Q3 and Q4 have a possible 0-value response in case none of the eHMI or iHMI is observed. It is intended to include a manipulation check that is answered before providing any specific value for the related question.

B. Direct Measurements

The distance to the pedestrian can be computed at any time during the experimentation using the vehicle positioning system. The speed of the vehicle can be directly read from the log files, and in combination with the distance, the Time To Collision (TTC) can be computed as TTC = d/v. These three variables can precisely describe the scene relative to the pedestrian's point of view. Figure 5 is an example of these three variables for an experiment with the optimal deceleration maneuver. Note that these temporal representations are meaningless without the proper labeling of the crossing event. All trajectories are identical for each type of braking maneuver and the only thing that changes is the interaction of each participant with the vehicle, i.e. the decision to cross.



Fig. 4. Crossing event example. Vehicle lane is defined by road marks unequivocally for all the experiments.



Fig. 5. Example of measured variables at the crossing event.

C. Crossing Decision Event

The crossing decision event is defined as the moment in which the pedestrian makes the decision to cross. We follow the hypothesis that the decision to cross is a hidden state that has an external and a posterior manifestation that can be observed. The "delay" between the decision and the external manifestation can vary depending on the person and the situation. Alternatively to the crossing decision event, we propose to use the crossing event as the metric to evaluate the behavior of the pedestrian using direct measurements in the experiments. The crossing event is defined as the frame in which the pedestrian enters the vehicle lane and physically exposes his/her body to a potential and real injury. The background idea is that an early crossing decision will produce an early crossing event and a late crossing decision will produce a late crossing event. The main difference is that the crossing event is not a hidden state. It is directly observable and can be unequivocally identified when the vehicle lane is defined using the lane markings. Figure 4 shows the vehicle lane boundary and the crossing event, defined as the moment in which the pedestrian enters the vehicle's lane. See figure 6 for a crossing sequence description example.

V. RESULTS

This section presents and analyzes the responses to the questionnaire and the measured variables for each type of interaction with the vehicle. For the analysis of the questionnaires, descriptive statistics and the Wilcoxon signed-rank test has been used. For the analysis of the direct measures, the Student t-test has been used to extract information.

A. Questionnaire Results

Questions from Q1 to Q6, are analyzed using the Wilcoxon Signed Rank test. Question 7 is a choice answer and its analysis is limited to the mode values.

The analysis of the participants' answers to questions from Q1 to Q6 will follow the same structure for each question. The alternative hypothesis matrix has been selected as the way to express categorical statements comparing the answers obtained after each experiment. As the null hypothesis, we propose $H_0 : \mu_i \leq \mu_j$ and as the alternative hypothesis, we take $H_1 : \mu_i > \mu_j$. A check-mark in a specific cell in table II means that H_0 is rejected and H_1 is accepted when comparing the answers provided in test *i* (left column) with test *j* (top row). In this specific context, the rejection of H_0 means that there is a difference with statistical significance between the answers for test *i* and *j*, and the answers for test *i* have a higher score in the Likert scale than for test *j*. The significance level has been selected to α =0.05 for all the probes. Experiment numbers are defined in table I.

TABLE II WILCOXON SIGNED RANK TEST FOR QUESTIONS Q1-Q6



Table III summarizes as a frequency table the answers for question Q7. It can be observed that the preferred communication mode for the internal HMI is the combination of audio and video simultaneously.



to the crosswalk.



sees the vehicle approaching.

(a) Initial position of pedestrian back (b) The pedestrian turns and faces the (c) The pedestrian starts walking and (d) At this point, the pedestrian is hesitating to cross.



the vehicle's reaction.

(e) The pedestrian is still waiting for (f) The pedestrian does not feel com- (g) The Pedestrian decides to cross (h) The Pedestrian crosses the crossfortable crossing while the vehicle is when the vehicle is almost stopped. walk. moving

Fig. 6. Example of vehicle-pedestrian interaction - exterior camera.

TABLE III FREQUENCY TABLE FOR QUESTION 7 ANSWERS

	Test number						
Answer	0	1	2	3	4		
Nothing	32	32	32	1	1		
Audio	0	0	0	7	10		
Video	0	0	0	9	9		
Both	0	0	0	15	12		

B. Direct Measurements Results

The analysis of the direct measures has been performed following the same alternative hypothesis matrix and the Student-t test. All the experiment measures are compared with all of them to obtain in which of them statistical significance can be observed. All the variables are continuous, consequently, the student's t-test for paired samples has been used. The same null $H_0: \mu_i \leq \mu_j$ and alternative $H_1: \mu_i >$ μ_i hypothesis are proposed for the alternative hypothesis matrix. A-check mark a specific cell in table IV means that H_0 is rejected and H_1 when comparing the measures from test *i* (left columns) against *j* (top row). In this specific context, a check-mark means that the distance, speed or TTC in test i is higher than for test j with a significance level $\alpha = 0.05$. To remove possible outliers and errors from different sources the two highest and lowest values have been removed from the analysis for all the variables.

C. Results Discussion

Based on the results shown in table II we can make the following statements: The optimal braking maneuver contributes to increasing the pedestrian's confidence in the vehicle (Q1: t1 vs t2 and t3 vs t4). The eHMI contributes to increasing the pedestrian's confidence in the vehicle. (Q1: t3 vs t1 and t4 vs t2) Pedestrians perceived the aggressive braking maneuvers as "more aggressive" or "less conservative" than the gentle breaking maneuvers (Q2: t2 vs t1

TABLE IV STUDENT T-TEST FOR DISTANCE, SPEED, AND TTC AT THE CROSSING EVENT





Fig. 7. Boxplot for distance variable at the crossing event.





crosswalk.

and t4 vs t3). The optimal braking maneuver contributes to increasing the passenger's confidence in the vehicle (Q4: t1 vs t2 and t3 vs t4). The iHMI contributes to increasing the passenger's confidence in the vehicle when stopping in a crosswalk with an aggressive braking maneuver (Q4: t4 vs t2). Passengers perceived the aggressive braking maneuvers as "more aggressive" or "less conservative" than the gentle braking maneuvers (O5: t2 vs t1 and t4 vs t3). We cannot state that the iHMI contributes to increasing the passenger's confidence in the vehicle when stopping in a crosswalk with an optimal braking maneuver (Q4: t3 vs t1). Attending to the frequency values provided for Q7 we can state that passengers found the combined mode (audio plus video) most helpful to build confidence than the audio or video choices for the iHMI. Based on the results shown in table IV we can make the following statements: The optimal braking maneuver contributes to increasing the distance at the crossing event when crossing in the crosswalk(dist.: t1 vs t2 and t3 vs t4). The eHMI in combination with the optimal braking maneuver contributes to increasing the distance at the crossing event (dist.: t3 vs t1). We cannot state that the eHMI contributes to increasing the distance at the crossing event when the AV brakes with the aggressive braking maneuver (dist.: t4 vs t2). This suggests that subjectively pedestrians feel safer but objectively they are not and the perception of risk dominates their actual behavior. As an additional finding, there were no statistical differences between the responses and measures obtained from the participants regardless of their initial role played.

VI. CONCLUSIONS AND FUTURE WORK

Questionnaires and direct measurements have proven that the iHMI and eHMI in combination with "gentle" or early braking maneuvers help to build trust in the AV when interacting with a pedestrian in a crosswalk for both the pedestrian and the passenger. However, there is a relevant difference between conclusions derived from questionnaires and measured variables. When we compare the experiments with the aggressive braking maneuver pedestrians express more confidence when using the eHMI than when not using it. However, it does not result in an earlier crossing event and consequently in an earlier crossing decision.

As a future work, this study can be extended by including interactions in non-signalized crossing areas and can be also replicated in Virtual Reality (VR) to study the gap between real and virtual interactions.

ACKNOWLEDGMENT

This work was funded by the HUMAINT project, Directorate-General Joint Research Centre of the European Commission and by Research Grants PID2020-114924RB-I00 and PDC2021-121324-I00 (Spanish Ministry of Science and Innovation) and S2018/EMT-4362 SEGVAUTO 4.0-CM and CM/JIN/2021-005 (Community of Madrid).

DISCLAIMER

The views expressed in this article are purely those of the authors and may not, under any circumstances, be regarded as an official position of the European Commission.

REFERENCES

- M. Li, B. E. Holthausen, R. E. Stuck, and B. N. Walker, "No risk no trust: Investigating perceived risk in highly automated driving," in *Proceedings of the 11th AutomotiveUI Conference*. Association for Computing Machinery, 2019, p. 177–185.
- [2] D. Fernández Llorca and E. Gómez, "Trustworthy autonomous vehicles," EUR 30942 EN, Publications Office of the European Union, Luxembourg, vol. JRC127051, 2021.
- [3] H. Detjen, S. Faltaous, B. Pfleging, S. Geisler, and S. Schneegass, "How to increase automated vehicles' acceptance through invehicle interaction design: A review," *International Journal of Human-Computer Interaction*, vol. 37, no. 4, pp. 308–330, 2021.
- [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, 2020.
- [5] I. Parra, H. Corrales, N. Hernández, S. Vigre, D. Llorca, and M. A. Sotelo, "Performance analysis of vehicle-to-vehicle communications for critical tasks in autonomous driving," in 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019, pp. 195–200.
- [6] 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, 2020.
- [7] S. Martín, R. Izquierdo, I. García Daza, M. A. Sotelo, and D. Fernández Llorca, "Digital twin in virtual reality for human-vehicle interactions in the context of autonomous driving," in *IEEE Intelligent Transportation Systems Conference*, 2023.
- [8] T. Lagstrom and V. M. Lundgren, AVIP-autonomous vehicles interaction with pedestrians. Gothenborg, Sweden: M.S. thesis, Dept. Product-Prod. Develop., 2015.
- [9] I. Parra, R. Izquierdo, J. Alonso, A. García-Morcillo, D. Fernández-Llorca, and M. A. Sotelo, "The experience of drivertive-driverless cooperative vehicle-team in the 2016 gcdc," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 4, pp. 1322–1334.
- [10] C. Dijksterhuis, B. Lewis-Evans, B. Jelijs, D. de Waard, K. Brookhuis, and O. Tucha, "The impact of immediate or delayed feedback on driving behaviour in a simulated pay-as-you-drive system," *Accident Analysis & Prevention*, vol. 75, pp. 93–104, 2015.
- [11] A. Tom and M.-A. Granié, "Gender differences in pedestrian rule compliance and visual search at signalized and unsignalized crossroads," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1794–1801, 2011.
- [12] P. Bazilinskyy, D. Dodou, and J. De Winter, "Survey on ehmi concepts: The effect of text, color, and perspective," *Transportation research part F: traffic psychology and behaviour*, vol. 67, pp. 175–194, 2019.
- [13] C.-M. Chang, K. Toda, X. Gui, S. H. Seo, and T. Igarashi, "Can eyes on a car reduce traffic accidents?" in *Proceedings of the 14th international conference on automotive user interfaces and interactive vehicular applications*, 2022, pp. 349–359.
- [14] 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*, 2019, pp. 295–307.
- [15] B. Mason, S. Lakshmanan, P. McAuslan, M. Waung, and B. Jia, "Lighting a path for autonomous vehicle communication: the effect of light projection on the detection of reversing vehicles by older adult pedestrians," *International journal of environmental research and public health*, vol. 19, no. 22, p. 14700, 2022.
- [16] M. Lynch, "Social constructivism in science and technology studies," *Human Studies*, vol. 39, pp. 101–112, 2016.
- [17] O. Benderius, C. Berger, and V. M. 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, 2017.
- [18] P. K. Murali, M. Kaboli, and R. Dahiya, "Intelligent in-vehicle interaction technologies," *Advanced Intelligent Systems*, vol. 4, no. 2, p. 2100122, 2022.
- [19] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 2961–2969.
- [20] R. Izquierdo, C. Salinas, J. Alonso, I. Parra, D. Fernández-Llorca, and M. A. Sotelo, "Testing predictive automated driving systems: Lessons learned and future recommendations," *IEEE Intelligent Transportation Systems Magazine*, vol. 14, no. 6, pp. 77–93, 2022.
- [21] A. Joshi, S. Kale, S. Chandel, and D. K. Pal, "Likert scale: Explored and explained," *British journal of applied science & technology*, vol. 7, no. 4, p. 396, 2015.