Communication of Automated Vehicles and Pedestrian Groups: An Intercultural Study on Pedestrians’ Street Crossing Decisions

PHILIP JOISTEN, ZIYU LIU, NINA THEOBALD, ANDREAS WEBLER, and BETTINA ABENDROTH, Technical University of Darmstadt, Germany

Fig. 1. An automated vehicle with an "walking person" external human-machine interface is approaching a zebra crossing in the single pedestrian (left) and group of three pedestrians (right) scenario.

Implicit as well as explicit cues are means of communication in driver-pedestrian interaction. With the introduction of automated vehicles (AVs), drivers can engage in non-driving related activities which rise new challenges of communication between AVs and pedestrians. In this context, external human-machine interfaces (eHMIs) are seen as a key contribution in building pedestrians’ trust towards AVs by enabling communication between them. However, a research gap exists regarding the communication of AVs and pedestrian groups. In an intercultural study we investigated the impact of the variables eHMI concept and group size on pedestrians’ street crossing decisions regarding (1) willingness to cross and (2) trust in AVs. Therefore, German (N = 126) and Chinese (N = 79) participants took part in an online-based video study. The results showed that a “walking person” eHMI had more stable effects with respect to the dependent variables in comparison to a “smiling face” eHMI in both countries. No main effect of group size on a pedestrian’s willingness to cross or trust in AVs was found. Nevertheless, qualitative data indicated an effect of group size in pedestrian-AV communication processes. Our results therefore contribute to the investigation of communication between AVs and pedestrian groups.

CCS Concepts:
• Human-centered computing → Empirical studies in HCI. User studies.

Additional Key Words and Phrases: automated vehicle, pedestrian, pedestrian groups, external human-machine interface, intercultural study, street crossing decision

ACM Reference Format:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.
Manuscript submitted to ACM

Urheberrechtlich geschützt / In Copyright
1 INTRODUCTION AND RELATED WORK

Driving is a social phenomenon that requires interactions between all road users to maintain a good flow of traffic and to guarantee the safety of all involved [23, 35]. In common driver-pedestrian interaction in urban traffic, implicit cues such as braking [1] as well as the non-verbal communication such as eye contact, gestures and facial expressions can be used to negotiate the right of way and predict intentions [9, 31].

However, with the introduction of automated vehicles (AVs), conventional driver-pedestrian communication strategies will no longer be available. To realize the expected benefits of an increased traffic efficiency and a reduction of accident as well as injury rates in urban areas [11], new modes of communication between AVs and pedestrians will be necessary [28]. Therefore, external Human-Machine Interfaces (eHMIs) are widely investigated [5, 8] and their influence in increasing trust towards AVs has already been demonstrated [15, 18]. Among the communication strategies studied, textual, light-based and symbol-based eHMIs can be differentiated [2, 3].

Studies have shown that well-designed symbol-based eHMIs are unambiguous [13] and recognizable from a greater distance [26]. In particular, symbol-based eHMIs provide the advantage of language-independent communication [4]. Additionally, culture-specific effects of eHMIs have already been identified: While an eHMI could improve pedestrians’ recognition rate of AV intentions only in Germany and the U.S. [34], a polite communication strategy increased compliant pedestrian behaviour solely in China [20].

Besides the influence of cultural factors on the impact of eHMIs, social factors such as group size have been shown to affect pedestrian behaviour [25, 27]. According to the ‘safety in numbers’ phenomenon [10], the self-reported safety feeling and the expectation that a crossing intention will be detected increases with group size [32, 37]. Pedestrians are thus at risk of paying less attention to the surrounding traffic situation and relying too heavily on the presence and behaviour of other pedestrians [12, 32, 33].

Although several studies have examined the influence of group size on pedestrian behaviour in conventional vehicle-pedestrian interaction and evaluated eHMIs in single-pedestrian scenarios, there is little research on the effects of group size on the interaction with AVs and on the effects of different eHMI designs in multiple-pedestrian scenarios. Therefore, this intercultural study addressed the following research question (RQ): What is the impact of the variables eHMI and group size on pedestrians’ street crossing decisions regarding (1) crossing willingness and (2) trust in AVs in Germany and China?

2 METHOD

A user study with a 2 x 3 x 2 mixed design was conducted as a video-based online survey. Country was selected as between-subjects factor with two levels (Germany vs. China). As within-subjects factors each participant experienced six traffic scenarios. In the traffic scenarios three eHMI designs (symbol “walking person”, symbol “smiling face” vs. without symbol) and two group sizes (group of three pedestrians vs. single pedestrian) were presented.

2.1 Sample

A total of 282 complete questionnaires were received between February and March 2021. We excluded 77 cases because of residence outside Germany or China, constant response pattern, very short completion time (i.e., less than three minutes) and failure to answer control questions correctly. This resulted in 205 cases included in the data analysis. The sample comprised of N = 126 German participants (54 female, 70 male, 2 diverse) and N = 79 Chinese participants (50 female, 29 male). The average age was M = 25.45 (SD = 4.70, Min = 18, Max = 57) years for the German participants.
and M = 34.08 (SD = 10.47, Min = 18, Max = 55) years for the Chinese participants. 26 German (20.6%) and 17 Chinese (21.1%) respondents stated that they had experience with AVs or had already participated in an experiment involving an AV with an eHMI.

Participants were recruited through word-of-mouth, social media and mailing lists. The questionnaire was presented in three languages (German, Chinese, English) and participants were requested to complete the questionnaire online using a tablet, smartphone or computer.

2.2 Procedure

Having selected a language to complete the online questionnaire, participants were first introduced to the aim of the study and the SAE Level 5 AVs [16] (i.e., self-driving cars). Then demographic data was collected (e.g., age, gender, experience with AVs). Afterwards, the participants watched seven videos which were presented in a randomized order. The instruction for each video was: 'Please watch the video by pressing the 'play button'. You can play the video multiple times. Please imagine the following scenario: You are walking along the sidewalk. You have an important appointment in a coffee shop which lies on the opposite side of the street. You are under time stress and need to get there on time. As you prepare to cross at a zebra crossing, you find that an automated vehicle is approaching (no driver sits in the car)'. Immediately after playing each video, participants answered the questions regarding the measures and control questions described below. Before the end of the survey, qualitative data on eHMI designs, group size and street crossing decisions was collected. On average, the study lasted for 8.4 minutes for German participants and 9.5 minutes for Chinese participants. No compensation was granted for participation in the study.

2.3 Materials

Six videos were recorded in a VR environment (Unity VR) to present six traffic scenarios which varied accordingly to the variables eHMI design and pedestrian group size (3 x 2 design, see Supplements). In the videos an AV (i.e., passenger car) without a driver present is approaching a zebra crossing in a daytime urban environment. A two lane road with a lane width of 3 meters each is visualized. An eHMI is mounted on the bumper of the AV (see Figure 1). The videos are 5 seconds long and are presented from the viewpoint of a pedestrian standing at a zebra crossing. At the beginning of the videos, the AV is 43.5 meters away from the pedestrian’s perspective and drives with a constant speed of 30 km/h. After 0.8 seconds the AV begins to decelerate with a constant deceleration rate of 2 m/s² and comes to a full stop 4.5 meters before the zebra crossing at the end of each video. In the traffic scenarios where an eHMI symbol is presented the symbol appears as soon as the car begins to decelerate. In the traffic scenarios with multiple pedestrians waiting at the zebra crossing, two additional pedestrians are standing in the participant’s field of view. These dummy pedestrians move their heads towards the approaching AV but do not step on the road.

The eHMI designs used in this study were selected based on previous studies [6, 15, 19] and differ regarding the perspective of communication [4]. The “walking person” eHMI represents an egocentric design, the “smiling face” eHMI represents an allocentric design.

In order to avoid habitual answers of participants, an additional dummy video was recorded and included in the experiment. In this video, neither a zebra crossing nor additional pedestrians are present. While driving at a steady speed of 30 km/h, the AV shows no intention to stop and displays an eHMI without any symbol.
2.4 Measures
After each video, we measured crossing willingness as a binary variable (“In this situation, will you cross or not?”). For a positive answer (i.e., yes), we measured trust in AVs, perceived safety and decision confidence adapted from [29]. Only the results for crossing willingness and trust towards AVs are reported in this paper. For a negative crossing willingness (i.e., no), we measured participants’ rating of trust in AVs (“To what extent did you trust the automated vehicle in the video?”). Then, the eHMI designs were evaluated in terms of message clarity, comfortability and preference [6, 19] after each video. Results regarding the evaluation of the eHMI designs are not included in this paper. All questions were presented as single items using a 7-point-scale from “not at all” to “extremely”.

For a manipulation check, two control questions were added after each video. Participants were asked which eHMI design was present in the video and whether additional pedestrians were waiting at the zebra crossing.

3 RESULTS
In this section, descriptive and inferential statistics are reported. For nominal data the Cochran’s Q Test was conducted. The non-parametric Friedman-Test was used for ordinal data. Bonferroni corrections were used for all post-hoc tests. All tests were conducted with a significance level of 0.05 and for the German and Chinese sample separately.

3.1 Crossing Willingness
Findings regarding crossing willingness are reported in Figure 2. Crossing willingness differed significantly between the six traffic scenarios for the German ($\chi^2(5) = 83.95, p < .001$) and the Chinese sample ($\chi^2(5) = 46.45, p < .001$) respectively.

In the single pedestrian scenarios, both the “walking person” eHMI and the “smiling face” eHMI statistically significantly improved crossing willingness for both countries compared to the eHMI without a symbol ($p < .004$). No statistically significant difference was found between the two eHMI concepts “walking person” and “smiling face” in the single pedestrian scenario for the German ($p = .174$) and Chinese sample ($p = 1.000$).

For the group of three pedestrians scenarios, both symbol-based eHMI designs statistically significantly improved crossing willingness for the German sample compared to the eHMI without a symbol ($p < .009$). For the Chinese sample, only the “walking person” eHMI statistically significantly improved crossing willingness compared to the eHMI without a symbol ($p = .003$), whereas the “smiling face” eHMI did not ($p = .513$). A statistically significant difference between the two symbol-based eHMI designs was found for the German sample ($p = .020$), but not for the Chinese sample ($p = 1.000$).

There was no statistically significant effect of group size (single pedestrian vs. group of three pedestrians) on participants’ crossing willingness for the same eHMI designs in both countries.

3.2 Trust in AVs
Findings regarding trust in AVs are reported in Figure 2. Between the six traffic scenarios, trust in AVs was significantly different for the German ($\chi^2(5) = 231.15, p < .001$) and the Chinese sample ($\chi^2(5) = 100.2, p < .001$) respectively.

For the German sample and the single pedestrian scenarios, both the “walking person” eHMI ($z = -2.1, p < .001$) and the “smiling face” eHMI ($z = -0.9, p < .001$) statistically significantly improved trust in AVs compared to the eHMI without a symbol. Additionally, German participants indicated a statistically significantly higher trust in AVs for the “walking person” eHMI compared to the “smiling face” eHMI ($z = 1.1, p < .001$). For the group of three pedestrians scenarios and compared to the eHMI without a symbol, both eHMIs statistically significantly contributed to participants’ trust in AVs.
in the German sample (“walking person” eHMI: $z = -2.2, p < .001$ ; “smiling face” eHMI: $z = -1.1, p < .001$). Again, higher trust in AVs was reported for the “walking person” eHMI when compared to the “smiling face” eHMI ($z = 1.1, p < .001$).

For the Chinese sample, both the “walking person” eHMI ($z = -2.0, p < .001$) and the “smiling face” eHMI ($z = -1.0, p = .017$) statistically significantly improved trust in AVs compared to the eHMI without a symbol in the single pedestrian scenarios. As for the German sample, the “walking person” eHMI resulted in higher trust in AVs compared to the “smiling face” eHMI for the Chinese sample ($z = 1.0, p = .008$). In the scenario with three pedestrians, only the “walking person” eHMI showed a statistically significantly effect compared to the eHMI without a symbol ($z = -1.4, p < .001$).

There was no statistically significant effect of group size (single pedestrian vs. group of three pedestrians) on participants’ ratings of trust in AVs for the same eHMI designs in both countries.

4 DISCUSSION

This intercultural study investigated the effects of AVs communicating with different symbol-based eHMI designs in a street crossing task. Additionally, the study examined the effects of pedestrian group size on crossing willingness and trust in AVs. Therefore, a video-based online survey with German and Chinese participants revealed main effects of eHMI design and country but did not show a statistically significant main effect of group size.

4.1 Effects of eHMI design, country and group size

The results of eHMI design showed that the “walking person” symbol improved crossing willingness and trust in AVs for both countries in comparison to an eHMI without a symbol. Hence, our results support prior work which showed positive effects of the ‘walking person’ design on pedestrians crossing decisions [13, 15]. Further, for the ‘smiling face’ design, trust in AVs was significantly and consistently lower compared to the “walking man” eHMI in the German
An egocentric perspective of communication ("walking person" eHMI) might be more beneficial for trust in AVs compared to an allocentric perspective [4], at least in Germany.

Chinese respondents could not benefit from the "smiling face" eHMI in the same way German respondents did. Since traffic rules and habits as well as traffic-related communication strategies vary across countries [14, 24], the same reactions to the same symbols cannot be expected everywhere. Additionally, this finding stresses the need to consider familiarity, meaningfulness, correctness and semantic distance [22] for eHMI design and evaluation. Still, cultural differences might diminish by a thorough introduction process to support comprehensibility of eHMI signals [30].

The study results indicated no significant main effect of group size on a pedestrian’s willingness to cross the street or trust in AVs. Nevertheless, participants stated that waiting in a group to cross the street increased their "feeling of safety" (P67) and "confidence to make crossing decisions" (P81). Some respondents believed that the AV could easier detect a group of pedestrians than single pedestrians ("The probability of detecting pedestrians would be higher", P123).

We assume that the effect of pedestrian group size depends not only on the number of group members but also on strength (i.e., individual factors that make a person influential) and immediacy (i.e., closeness in space or time) as proposed by [21]. Additionally, the waiting pedestrians in this study were just present and did not behave in a specific way (e.g., they did not step on the street to cross). No imitable (compliant or non-compliant) behaviour was shown, which has been demonstrated to have an impact on pedestrians’ decision-making when crossing streets [17, 36].

4.2 Limitations and Outlook

This study has some limitations as the scope of this study was limited by the means of a video-based online survey. First, in this study we investigated only one traffic scenario environment (i.e., a yielding AV at a zebra crossing) which could have led to habitual response behaviours, especially regarding crossing willingness. Second, we limited the eHMI designs on two symbols and examined them in comparison to an eHMI without a symbol. Other communication strategies (i.e., textual or light-based eHMIs) [5, 7] were not included in this study. It needs to be further discussed which communication strategy is best applicable for a targeted and distributed communication with single pedestrians and pedestrian groups. Therefore, more insights into the effects of AV communication strategies and pedestrians group sizes in various traffic environments and cultures are needed in order to derive recommendations for eHMI design.

5 CONCLUSIONS

This empirical study set out to explore the effects of group size on pedestrians’ crossing willingness and trust in AVs as well as the effects of different symbol-based eHMIs designs in a multiple-pedestrian street crossing scenario. The results were compared between two countries. Overall, a "walking person" eHMI showed more stable effects regarding the dependent variables than a "smiling face" eHMI design. No main effect of group size on pedestrians’ willingness to cross or trust in AVs was found. However, qualitative data indicated an effect of group size in pedestrian-AV communication processes. Our results therefore contribute to the investigation of communication between AVs and pedestrian groups. Still, a necessity remains to incorporate the group size and effects of social impact in further studies of AVs communication strategies.

ACKNOWLEDGMENTS

This research was funded by research project @CITY-AF, carried out at the request of the Federal Ministry for Economic Affairs and Energy (BMWi), under research project No. 19A18003M. The authors are solely responsible for the content.
REFERENCES


A SUPPLEMENTS

The video material used in this study is online at YouTube (https://youtu.be/7F-9vZRkExM) and can be downloaded from this URL: https://hessenbox.tu-darmstadt.de/public?folderID=MjlYdHJqeVM5a3V0MldLa0xqTGlF