Communication Between Automated Vehicles and Drivers in Manual Driving Vehicles: Using a Robotic Arms to Produce Gestures

Wei Zhang[1], Changxu Wu[1], Xianwen You[1], Leo Kust[2], Yijie Chen[3], Jinlei Shi[4]

[1] Department of Industrial Engineering, Tsinghua University, Beijing, China

[2] State University of New York at Buffalo, New York, the United States

[3] Arizona State University, Phoenix, the United States

[4] College of Computer Science and Technology, Zhejiang University, Hangzhou, China

Corresponding author: Changxu Wu, changxu.wu@gmail.com, Department of Industrial Engineering, Tsinghua University, Beijing, China

Communication Between Automated Vehicles and Drivers in Manual Driving Vehicles: Using a Robotic Arms to Produce Gestures

Abstract

Effective communication between automated vehicles and human drivers in manual driving vehicles is of great importance for traffic safety during the transition phase of automated vehicles. Gestures, which were widely used in road users' communication, were promising in conveying the intentions of automated vehicles naturally and intuitively without extra learning costs. However, the effect of gestures in conveying the automated vehicles' intentions on human understanding remains unknown. This study proposed the idea of adopting robotic arms to produce gestures. An experiment based on video recordings was conducted to explore the effect of arms type (slow-waved robotic arm (80 beats/min) *vs.* fast waved robotic arm (120 beats/min) *vs.* human arm) and gesture type (taking the road *vs.* giving the road) on the participants' objective responses and subjective opinions. A total of 30 participants were recruited as human drivers in a manual driving vehicle, who received and responded to the gestures transferred by an encountering automated vehicle. Results indicated that regardless of the gesture type, the slow-waved robotic arm led to a longer response time (mean ± SD: 4.871 ± 0.947 s) and lower response accuracy (88.3 ± 32.4 %) when compared with the human arm (response time: 4.457 ± 0.727 s, response accuracy: 95.0 ± 22.0%). It was also rated less understandable and comfortable than the human arm. Nevertheless, the fast-waved robotic arm not only exerted as fast (4.484 ± 0.818 s) and accurate responses (98.3 ± 12.9 %) as the human arm but was also rated as understandable, polite, and comfortable as the human arm. This indicated the implication of conveying gestures by utilizing the fast-waved robotic arm (120 beats/min) to facilitate effective communication from automated vehicles to human drivers in manual driving vehicles. The present study's findings provided reference implications for manufacturers and designers to adopt this gesture-based communication method to develop safe and user-friendly automated vehicles.

Keywords: automated vehicle; human drivers; gestures; robotic arms

1 Introduction

1.1 Background

The future transportation system would be largely changed with the development of automated vehicles, which have great potential in optimizing traffic flow, reducing road accidents, alleviating air pollution, improving driver comfort, and so on (Anderson et al., 2014; Aoki et al., 2021; Maurer et al., 2016; Millard-Ball, 2018). The Society of Automotive Engineers (SAE) classified driving automation into six levels, with the automation above level 3 no longer requiring drivers to supervise the traffic environment (SAE International, 2021). A long transition phase would emerge before manual driving vehicles were completely replaced by automated vehicles, which contained both manual driving and automated vehicles on the road (Aoki et al., 2021). Thus, under this transition phase, the automated systems and human drivers from manual driving vehicles must understand each other and cooperate to maintain a safe and orderly traffic environment (Aoki et al., 2021; Xing et al., 2021).

From the automated vehicles' perspective, algorithms can be designed and improved constantly to understand the human drivers' behaviors by processing the real-time information gathered by lidar, radar, and cameras (Deo et al., 2016; Flores et al., 2018; Mukhtar et al., 2015). However, humans processed the road information with their brains, which had limited cognitive resources (Kahneman, 1973). Their understanding of the traffic environment usually relies on mutual communication such as the direct physical movements of the vehicle (e.g., yielding or approaching the left lane), the signals sent by the drivers (e.g., honking and lighting), and non-verbal communication methods, which included facial expressions, eye contact, gestures (e.g., waving hands), body movements, and the voice and tone of speech (Merten, 1997; Sucha et al., 2017). However, in the mix of manual driving and automated vehicles, human drivers cannot always communicate directly and timely. It is because the driver in automated vehicles (especially for automation above level 3) are more likely to engage in non-driving related tasks without participating in driving tasks (Maurer et al., 2016). Inappropriate actions may be taken if manual drivers misunderstand the intentions of automated vehicles, which would increase accident risks and impact traffic efficiency (Fuest et al., 2020). Therefore, it is important for the automated system to effectively and properly communicate with human drivers to make them understand its intentions (Lee et al., 2010; Rasouli & Tsotsos, 2019).

1.2 Gestures as promising communication methods

In the past few decades, researchers have widely explored effective communication methods and messages between automated vehicles and road users (Xing et al., 2021). The communication methods included passive signs such as delivering alerts through bracelets or smartphones (Mahadevan et al., 2018; Rahimian et al., 2016), vehicle motion characteristics (Fuest et al., 2020; Rettenmaier et al., 2021), and external human-machine interfaces (eHMIs) such as transferring messages by LED light panels with textual or specific signs (Clamann et al., 2017; De Clercq et al., 2019; Faas & Baumann, 2021; Fiedler & Spelten, 2017; Habibovic et al., 2018). Moreover, the communication messages transferred by those methods can be classified into base messages (i.e., the driving modes and perception of automated vehicles), intention messages (i.e., inform the automated vehicles' intention and decision for the next action), and instruction messages (i.e., instruct the road users how to act) (Xing et al., 2021; Zandi et al., 2020; Zhang et al., 2017), with the intention messages was relatively more important than others.

However, several limitations existed in the communications above. First, extra learning costs existed in some designed eHMIs. For instance, the participants needed a pretraining to learn the meaning of the proposed light interface that represented the states of automated vehicles before making a judgment (Habibovic et al., 2018). Second, some proposed interfaces may not be universally applicable or equally necessary when communicating between automated vehicles and people from different backgrounds. For instance, based on an international study, Zandi et al. (2020) found that people from different countries showed different views on the importance of communication messages between automated vehicles and pedestrians. Take the message "Warning, I am dangerous!" as an example, it was served quite important for people in the USA, while it was rated close to meaningless for people in Germany in a circumstance when automated vehicles ignored the pedestrian's intentions (Zan di et al., 2020). It might be because of the variety of languages, road cultures, and norms, people from various regions interact with automated vehicles differently (Currano et al., 2018; Färber, 2016; Li et al., 2020), and their expectations for human-machine interfaces varied (Alexander et al., 2017; Weber et al., 2019).

By contrast, gestures, which had been widely used in daily life, served as a pervasive and promising communication method (Maurer et al., 2016). Compared with previously mentioned communication methods, conveying intentions through gestures has unique advantages. First, gestures are generally natural and intuitive (Khan & Tunçer, 2017). Humans can perceive and understand gestures immediately and directly at first glance with little extra learning cost. Second, the understanding of gestures may be relatively more consistent across people from various backgrounds when compared with eHMIs, especially for the text and speech content that varies by language. Although some symbol-based eHMIs were proved in uniform intention recognitions (Singer et al., 2020) and stable effects on crossing decisions (Joisten et al., 2021) of humans from different nationalities, some of these effects might be attributed to the pre-explanation of symbols in the study (Singer et al., 2020). Based on the findings of gestures consistency across countries (Meier et al., 2014; Pika et al., 2009; Gupta et al., 2016), conveying intentions of automated vehicles by gestures has the advantage of being understood without pre-explanations by people from different countries. Third, gestures can convey comprehensive and clear signals by simply waving hands or arms. For instance, a gesture of offering with the palm facing upward denoted "Go ahead, I am giving the road for you" (Färber, 2016; Maurer et al., 2016). Hence, conveying the intentions of automated vehicles in gestures would be a humanoid way, which may prompt the direct and rapid human understanding of the automated vehicle's intentions by serving it as a driver in the traditional traffic environment.

1.3 Convey gestures by robotic arms

With the potential of gestures in conveying the intentions of automated vehicles, we raised the following question: how to convey these gestures by automated vehicles? Previous studies indicated that programming gestures into robotic arms could effectively prompt human-machine/computer interactions (Gleeson et al., 2013; Salem et al., 2012; Sheikholeslami et al., 2015, 2017). For example, in exploring the efficacy of the robot hand in expressing instructional gestures for human-robot interactions in assembly tasks, Sheikholeslami et al. (2017) found that humans can recognize the robot gestures with a recognition rate greater than 60%. Gleeson et al. (2013) also observed that humans could easily interpret the gestures transferred by the robotic system. Apart from being recognizable, the robot gestures can also gain positive evaluation from humans in the quality of presentation and perception (Salem et al., 2012). Given the effectiveness of robotic gestures in human-robot interaction, transferring gestures by robotic arms may be a promising way of realizing the communication between automated vehicles and manual drivers.

Some critical parameters, which may influence the effectiveness of robotic arms in transferring gestures (e.g., speed, amplitude, repetition, and arm extension) (Deshmukh, Craenen, Foster, et al., 2018; Xu et al., 2013), should be considered when adopting the robotic arms. Particularly, the speed of gestures was one of the major parameters that characterized any natural and synthetic gestures (Deshmukh, Craenen, Vinciarelli, et al., 2018). Researchers indicated that speed was an essential parameter that influenced the human perception of robotic gestures (Berger et al., 2021; Moon et al., 2013; Riek et al., 2010) and was an important cue that humans used to interpret the robot's intentions (Lohse et al., 2013). For example, Berger et al. (2021) found that when changing the speed of gestures, humans perceived the gestures with different meanings during the interaction with robots. The diversity of gestures for changing the speed would be useful in social human-machine interaction. However, when it comes to human drivers and automated vehicle communication, the most important thing is to ensure that human drivers can correctly perceive and understand the meaning of gestures in automated vehicles. If drivers misunderstand the meaning of gestures and misinterpret them when the waving speed of robotic arms changes, unpredictable traffic accidents may happen. Therefore, exploring whether human drivers can correctly understand the meaning of robotic gestures at different speeds is crucial. However, to our knowledge, no studies had applied the robotic arms in conveying gestures of automated vehicles to road users, leaving the effect of robotic arms with different speeds on human drivers and automated vehicle communications remain unknown.

Based on the summary of the studies above, gestures produced by robotic arms have great potential in effectively conveying the intentions of automated vehicles. Nevertheless, utilizing the robotic arms to produce gestures in communication between the automated vehicles and human drivers in manual driving vehicles was unexplored. Some researchers utilized gestures in human-automated vehicle communications. For instance, Epke et al. (2021) investigated the gestures in communication between road users and automated vehicles. However, they focused on the hand gestures used by pedestrians in intention communication to the automated vehicle rather than using the gestures to indicate the automated vehicle's intentions. Oudshoorn et al. (2021) designed a gesture-based eHMI, which consisted of flaps on the left, right, and top of the automated vehicles to show the yielding or non-yielding intentions to road users. Nevertheless, the gestures conveyed by flaps were inspired by the elephants, which differed from the humanoid gestures based on robotic arms that the present study suggested. Moreover, when most studies focused on the communication between the automated vehicles with pedestrians (Epke et al., 2021; Fuest et al., 2020; Oudshoorn et al., 2021; Rettenmaier et al., 2021), the communication between the automated vehicles with human drivers in manual driving vehicles was less concerned.

1.4 Aim of the present study

This study aimed to (1) investigate the effect of using robotic arms to produce gestures in conveying the intentions of automated vehicles compared with the human arm; (2) explore whether the speed of robotic arms would influence the human responses and understandings of different gestures.

In this study, the automation above level 3 was considered for the reason that the drivers were no longer required to monitor the road (SAE International, 2021) and were more likely to be absent in intention communications with other road users. In addition, two basic gestures when it comes to the conflicts of using roads were considered: taking the road (i.e., passing the road directly) and giving the road (i.e., waiting for others to pass) (Weber et al., 2019; Gupta et al., 2016).

2 Methods

2.1 Participants

A total of 30 participants (15 males and 15 females) ranging from 18 to 30 (Mean = 21.43, SD = 3.18) engaged in the experiment. All participants were required to have the normal or corrected-to-normal vision and own valid driving licenses. Their driving experience was from one to five years (Mean = 2.0, SD = 0.83). The present study was approved by the Institutional Review Boards (IRB) of University at Buffalo.

2.2 Experiment Design and measures

This study adopted a 2 × 3 within-subject design. Independent variables were the gesture type (taking the road *vs.* giving the road) and arm type (human arm *vs.* fast-waved robot arm *vs.* slow-waved robot arm). Two commonly used gestures were transferred to participants. A holding vertical hand gesture was designed as a "taking the road", and a moving gesture by the palm facing toward them was designed as a "giving the road". Three types of arms were utilized to transfer these gestures. That is, participants experienced two robotic arms with fast (120 beats per minute) and slow (80 beats per minute) speeds and one human arm (set as the control condition). A total of six conditions were presented to participants.

This study collected the participants' objective responses and subjective opinions about gestures conveyed by automated vehicles. Two measurements were recorded to evaluate the participants' behaviors: 1) Response time: the time between the onset of gestures and the moment the participants pressed the spacebar on a keyboard, which action meant they understood the gestures. 2) Response accuracy: the participants responding accurately (recorded as 1) or not (recorded as 0) for corresponding transferred gestures. Three measurements were collected to evaluate the participants' subjective attitude toward the gestures conveyed by the automated vehicle: 1) Understanding: the understanding of the participants for gestures transferred by different arms by asking "how well did you understand the gesture given?"; 2) Politeness: the politeness of gestures transferred by different arms by asking "how polite was the gesture given?"; 3) Comfort: the comfort level the participants experienced for gestures transferred by different arms by asking "how comfortable did you feel by the gesture given?". To minimize the misuse of a midpoint and to find an optimal response, the above subjective measurements were all based on a four-point Likert scale (omit the midpoint) from "not at all" to "very much" (Chyung et al., 2017).

2.3 Materials

For safety considerations, this study adopted the laboratory experiment. The real gesture transferring scenarios were filmed into videos in advance. The following were steps to generate the gesture transferring scenarios.

Step1: building a robotic arm and generating gestures

The robotic arm included the plastic hand and arm parts. In building the plastic hand, the sizing was based on average male hand anthropometry, with the hand length being 0.188 meters and the palm's width being 0.08 meters (Pheasant & Haslegrave, 2018). The color of the hand and arm were chosen for visual salience compared with the rest of the car’s dashboard. To create a visual distinction between the hand and the arm, we colored the hand portion white and the arm portion yellow. The hand was operated by positioning the wrist at an angle with the thumb facing upward and oscillating from the wrist.

Two motions were designed to consider “taking the road” and “giving the road” gestures. The “taking the road” gesture originated from the wrist with the palm facing downward and raised to a 90-degree angle with the artificial palm perpendicular to the dash (see Figure 1. c), which indicated that the automated vehicle wanted the participants to wait. The “giving the road” gesture included horizontal-waved hand and arm motions (see Figure 1. d). In this way, participants received the signal with the intention of the automated vehicle wanting them to move first. The waving speeds of fast and slow arms were 120 beats per minute and 80 beats per minute. An apparatus that triggered the arms was set on the dash to the right of the driver’s seat and located 0.2 meters away from the steering wheel. In this way, the participants can easily see the gestures on the automated vehicles.

Step2: filming and editing videos

A total of six gesture-transferred scenarios which corresponded to six experimental conditions were filmed into videos at a four-way crossroad at the University at Buffalo’s campus. We invited a middle-aged male driver to attend the filming. The driver sat in a car to simulate the driver in the automated vehicle. The car’s left turn signal was on and flashing to simulate a scenario in which the car intended to make a left turn. In the human arm condition, gestures were transferred by the driver's hand and arm when approaching the crossroad. The driver faced the camera with sunglasses and looked downward, wherein eye contact and facial expressions would be minimized (see Figure 1. a, b). In the robotic arm conditions, gestures were transferred by the fast-waved and the slow-waved robotic arms, respectively. The driver lowered his head to simulate engaging in a non-driving related task. At the same location where the driver started waving his hand, the apparatus triggered the fast-waved robotic arm or the slow-waved robotic arm to transfer two gestures.

After recording, each video was edited to be six seconds long, with two seconds of lead time before any action was taken by either the robotic arm apparatus or the driver and four seconds for the gesture to be viewed. Next, the videos were coded into a Visual Basic Application (VBA) format, which allowed the videos to be played and the responses of the participants to be documented.

2.4 Experiment apparatus

The experiment was conducted in a simulation room equipped with a keyboard, a 60-inch TV screen, a driving set with an adjustable driving seat, a Logitech Driving Force steering wheel, a throttle pedal, and a brake pedal (Logitech Inc., Fremont, CA).

Participants sat in the driving seat in front of the TV screen to simulate the drivers who would encounter an automated vehicle. The distances between participants and the screen were determined by the screen size and video. The vehicle shown in the video would be perceived as the same size and distance from the participant as in the real world, with the visual angle also relatively fixed based on the data collected along with the video.

2.5 Procedure

First, participants were welcomed. They were instructed to sign a consent form and a questionnaire related to their age, gender, driving experience, etc. Next, participants were instructed to sit in the driving seat and imagine that they were a driver manually driving approaching an uncontrolled crossroad to go straight. At the same time, participants were told that an oncoming automated vehicle would turn left; the vehicle automation was above level 3, in which the driver in the automated vehicles would not participant in driving in most cases. Participants were told that the path of the front vehicle would intersect with their driving path; then, they should therefore know the intention of the oncoming automated vehicle by seeing gestures, which would be given by the driver or a robotic arm in the place of the driver. The experimenter checked and confirmed the participants' correct understanding of the experimental task before the experiment. In the experiment, the previously edited videos corresponding to experimental conditions were presented to participants. The participants pressed the spacebar immediately once they understood the meaning of the gesture and reported its meaning transferred. At the end of each trial, the participants completed a four-point Likert scale to report their understanding of the received gesture and evaluate the politeness and comfort level of the gesture transferred from the automated vehicle. Participants were not given feedback on whether their understanding of the gesture was correct. A total of six trials corresponding to six scenario videos were included in the experiment. The order of six videos presented to participants was counterbalanced by the Latin Square to avoid the sequential effect (Bradley, 1958; Kantowitz et al., 2014) (see more details in Appendix Table A1). After finishing the experiment, each participant was thanked and compensated 10 dollars. The whole experiment lasted approximately 30 minutes.

2.6 Data analysis

The boxplots of response time, understanding, politeness, and comfort ratings were plotted by MATLAB R2021b (MathWorks, Inc. MA). Outliers pinpointed by the boxplot were excluded in the data analysis. Then statistical data were analyzed using IBM SPSS 26.0 (IBM, Inc). Firstly, the normality distribution of the response time was checked by the one-sample Kolmogorov Smirnov test. The original data of response time had positive skewness; thus, we transformed the data by the Box-Cox method before analysis (Box & Cox, 1964). Then a "mixed model approach," which can consider the fixed and random effects, was used (Baayen et al., 2008; Breslow & Clayton, 1993; Wan & Sarter, 2022). A linear mixed model (LMM) was adopted for analyzing the response time, and a generalized linear mixed model (GLMM) with a logit link function was conducted for binary variables (i.e., response accurate or not whereby 1 = accurate responses and 0 = inaccurate responses). In this study, the gesture type, arm type, and interactions were considered fixed variables, while the individual variance was considered the random variable. We used the paired t-test for the post-hoc test if the main effect was significant and the simple effect analysis if the interaction effect was significant. The Friedman test was first conducted for subjective ratings, which is a non-parametric statistical tool for ordinal data with repeated measures (Friedman, 1937). A Wilcoxon signed-rank test was employed for the post-hoc test if the Friedman test showed statistically significant differences (Wilcoxon, 1992). Due to the limitation of the Friedman test in exploring the interaction effect, we analyzed the effect of arm type on subjective ratings in two gesture conditions, respectively. In addition, all the significance level of the analyses was set to 0.05 and was corrected by the Bonferroni adjustment.

3 Results

3.1 Objective responses

Response time

Firstly, a boxplot with median and interquartile range (IQR) of response time was presented in Figure 2 (a). From the boxplot, it can be noted that the median value of response time of the slow-waved robotic arm was slightly higher than that of the fast-waved robotic arm and the human arm, regardless of gesture type. Moreover, from the whiskers, the range of response time in the slow-waved robotic arm condition was relatively higher than that in the human arm and the fast-waved robotic arms. This indicated that participants had a relatively less consistent response time to the slow-waved robotic arm.

The LMM analysis indicated the significant main effect of arm types on response time ($F_{(2, 154)} = 9.443$, $p < 0.001$). The post hoc test revealed that participants reacted faster under the human arm (Mean = 4.484, SD = 0.818) and the fast-waved robotic arm (Mean = 4.457, SD = 0.727) conditions than under the slow-waved robotic arm condition (Mean = 4.871, SD = 0.947) ($p < 0.001$ for each comparison). However, there were no significant differences between the human arm and the fast-waved robotic arm in response time. No main effect of gesture type ($F_{(1, 154)} = 0.582$, $p = 0.447$) nor the interaction effect of arm type and gesture type ($F_{(2, 154)} = 0.240$, $p = 0.787$) on response time were observed. The detailed descriptive statistical values for response time in terms of arm type and gesture type were listed in Table 1.

Response accuracy

The detailed descriptive statistical values for response accuracy were listed in Table 1. It can be noted that the mean response accuracy was highest in the human arm condition (giving the road: 100 ± 0 %; taking the road: 96.7 ± 18.3 %), then the second high in the fast-waved robotic arm condition (giving the road: 93.3 ± 25.4 %; taking the road: 96.7 ± 18.3 %), while the mean response accuracy of the slow-waved robotic arm (giving the road: 90 ± 30.5 %; taking the road: 86.7 ± 34.6 %) was lowest with largest standard deviation among three types of arms in each gesture condition.

The GLMM analysis showed that the main effect of arms type ($F_{(2, 174)} = 10.235$, $p < 0.001$) on response accuracy was significant. The mean response accuracy of the human arm, the fast-waved robotic arm, and the slow-waved robotic arm were 98.3% (SD = 12.9), 95.0% (SD = 22.0), and 88.3% (SD = 32.4), respectively. The post-hoc test suggested that the drivers responded to the human arm more accurately than the slow-waved robotic arm ($p < 0.05$). The differences between the fast-robotic and the human arm (p = 0.266) and the slow-waved robotic arm (p = 0.266) were not significant. There was an interaction effect of the arm type and gesture type on response accuracy ($F_{(2, 174)} = 3.688$, $p < 0.05$). However, the simple effect analysis showed no significant differences for each paired comparison. Moreover, the main effect of the gesture type on response accuracy was also not significant ($F_{(1, 174)} = 2.408$, $p = 0.123$).

3.2 Subjective ratings

Statistical descriptions, including the median, IQR, and mean rank of three subjective ratings, were listed in Table 2, and the data visualization was presented in Figure 2 (b) (c) (d).

Understanding for gestures

For giving the road gesture, a Friedman test revealed a significant effect of arm type on understanding for gestures ($\chi^2(2) = 8.000$, $p < .05$). The post-hoc test showed the understanding ratings of the slow-waved robotic arm (median = 4.0, IQR = 3.0 –

4.0) was significantly lower than the human arm (median = 4.0, IQR = 4.0 – 4.0) (Z = – 2.310, $p < .05$, r = – 0.42) (see Figure 2 (b) and Table 3). The differences between the fast-waved robotic arm (median = 4.0, IQR = 3.0 – 4.0) with the human arm as well as it with the slow-waved robotic arm on understanding were not significant.

For taking the road gesture, there was also a significant effect of arm type on understanding for gestures according to the Friedman test ($\chi^2(2) = 8.000$, $p < .05$). The post-hoc test suggested that the understanding ratings for gesture of the slow-waved robotic arm (median = 3.0, IQR = 2.0 – 3.25) was significantly lower than the human arm (median = 3.0, IQR = 3.0 – 4.0) (Z = -2.556, $p < .05$, r = - 0.47) and the fast-waved robotic arm (median = 3.0, IQR = 3.0 – 4.0) (Z = - 2.982, $p < .01$, r = - 0.54) (see Figure 2 (b) and Table 3). The differences between the fast-waved robotic arm and the human arm on understanding was not significant.

Perceived politeness for gesture

The Friedman tests suggested no significant differences of arm type on politeness for either giving the road gesture ($\chi^2(2) = 0.175$, $p = 0.916$) or giving the road gesture ($\chi^2(2) = 0.485$, $p = 0.785$).

Comfort ratings for gestures

For giving the road gesture, a Friedman test suggested the significant differences of arm type on comfort ratings for gestures ($\chi^2(2) = 7.000$, $p < .05$). The post-hoc test showed the significantly lower comfort ratings of the slow-waved robotic arm (median = 3.0, IQR = 2.0 – 4.0) than that of the human arm (median = 3.0, IQR = 3.0 – 4.0) (Z = – 2.500, $p < .05$, r = – 0.46) and the fast-waved robotic arm (median = 3.0, IQR = 3.0 – 4.0) (Z = – 2.066, $p < .05$, r = – 0.38) (see Figure 2 (d) and Table 3). No significant differences between the fast-waved robotic arm and the human arm on understanding was observed.

For taking the road gesture, there was also a significant difference of arm type on comfort ratings for gestures ($\chi^2(2) = 26.079$, $p < .001$). The post-hoc test suggested the significantly lower comfort ratings of the slow-waved robotic arm (median = 2.0, IQR = 2.0 – 3.0) than that of the human arm (median = 3.0, IQR = 2.0 – 4.0) (Z = – 4.234, $p < .001$, r = - 0.77) and the fast-waved robotic arm (median = 3.0, IQR = 2.0 – 3.25) (Z = -3.578, $p < .001$, r = – 0.65) (see Figure 2 (d) and Table 3). Again, the differences between the fast-waved robotic arm and the human arm on understanding was not significant.

4 Discussion

Within the mixed environment of manual driving vehicles and automated vehicles, it is important for automated systems to communicate with human drivers in manual driving vehicles effectively. In exploring whether gestures produced by robotic arms can be utilized to transfer the automated vehicles' intentions to human drivers, this study investigated the effects of various types of arms (human arm, fast-waved robotic arm, and slow-waved robotic arm) and gestures ("giving the road" and "taking the road") on drivers' objective responses and subjective opinions.

4.1 The effect of robotic arms on drivers' responses

In terms of the effect of using robotic arms in transferring gestures on the participants' responses, results suggested that the fast-waved robotic arm had a similar effect to the human arm in drivers' safe responses and positive subjective opinions. Specifically, regarding response time, drivers averagely responded to the fast-waved robotic arm (4.457 ± 0.727 s) as quickly as the human arm (4.484 ± 0.818 s). Moreover, they responded to both above arms more swiftly than the slow-waved robotic arm (4.871 ± 0.947 s) from the onset of gestures. For one thing, this response time was relatively shorter than that in the previous study (Fuest et al., 2020), which found the pedestrians' intention recognition time to the intention of the automated system was between 4.7 s – 6.7 s based on a video experiment. It might lie in the differences in intention communication objects and methods between the present study and Fuest et al. (2020)' s. The present study focused on the communication of automated vehicles to human drivers, while Fuest et al. (2020) focused on pedestrians. Pedestrians were more vulnerable than drivers and may feel more uncertain about intentions they understood; thus, they needed more time to press the button. Moreover, the present study utilized gesture-based arms for communicating between automated vehicles and humans. However, the intentions of automated vehicles were conveyed by changing speed in Fuest et al. (2020)' study, the latter of which may take the participants some time to recognize the speed variance of the automated vehicle before interpreting its intention. For another, the response time differences were 0.387 s between the slow-waved robotic and human arm and 0.414 s between the slow-waved robotic and the fast-waved robotic arm. These time differences were close to that in a previous study (Sullivan et al., 2008), which found that the differences in drivers' response time to warning systems with two different reliability was about 0.375 s. While the Sullivan et al. (2008)' s study focused on the traffic safety issue, the response delay in the present study might influence the traffic efficiency or stability in the mixed traffic environment (Do et al., 2019; Guériau & Dusparic, 2020; Kesting & Treiber, 2008; Mahmud et al., 2017). Based on the finding that the increase of drivers' reaction time and acceleration led to a decrease in the traffic stability (Kesting & Treiber, 2008; May, 1990), it could be inferred that the cumulative effects of many pairs of vehicle-to-drivers communication with the extra response time would reduce traffic stability and may give a burden on the traffic congestions, especially in urban areas. Nevertheless, the present study did not explain how the response time differences of different arms would affect the traffic efficiency or stability, which were supposed to be explored in future studies.

Moreover, the response accuracy showed a similar trend for different arm types. The fast-waved robotic arm led to a similarly high response accuracy (98.3 ± 12.9 %) as the human arm (95.0 ± 22.0 %). Furthermore, both above arms resulted in significantly more accurate responses than the slow-waved robotic arm (88.3 ± 32.4%). Although the mean response accuracy of different arms did not differ much (the slow-waved robotic arm can achieve 88.3 % on average), some subsequent effects should be considered. For instance, if drivers misunderstand the automated vehicles' attention and adopt the wrong actions (e.g., driving directly when the automated system tends to take the road or keeping waiting when the automated system gives the road to drivers), the risks of accidents and congestion would be raised. Therefore, the fast-waved robotic arms, which can lead to fast and accurate responses, were supposed to be considered primarily for application.

In summary, it can be inferred that response time and accuracy results supported the fast-waved robotic arm rather than the slow-waved robotic arm. On the one hand, it may be because the participants' recognition of gestures required a certain amplitude of motions of arms. As the amplitude of the slow-waved robotic arm changed relatively slowly, it was hardly distinguishable at first sight. Only after the motion reached a certain amplitude could the meaning of gestures be formed. With the increment of waving speed, the intention of gestures can be perceived and understood quickly. Thus, participants may spend less time recognizing and responding to the fast-waved robotic arm than the slow-waved robotic arm. On the other hand, fast robot motions were related to human beings' perceived arousal and valence during the human-robot interactions (Saerbeck & Bartneck, 2010; Zoghbi et al., 2009). It may be because the fast-waved robotic arms enhanced the drivers' arousal level, thus, leading to their faster responses and higher accuracy.

4.2 The effect of robotic arms on drivers' subjective opinions

The results of the subjective ratings were mostly consistent with that of the objective behaviors, which indicated that the fast-waved robotic arm was as favorable as the human arm while the slow-waved robotic arms were less favored. Specifically, although the median value of understanding for three arms was the same in giving the road (4.0 for all arms) or taking the road conditions (3.0 for all arms), the IQR indicated the differences. When the participants showed the most consistent and best understanding of the human arm ($IQR_{\text{giving the road}}$: 4.0 – 4.0, $IQR_{\text{taking the road}}$: 3.0 – 4.0), the understanding scores of the slow-waved robotic arm was rated with higher variance ($IQR_{\text{giving the road}}$: 3.0 – 4.0, $IQR_{\text{taking the road}}$: 2.0 – 3.25). This suggested that participants were more uncertain about whether they correctly understood the intentions when gestures were transferred by the slow-waved robotic arms. Based on the statistical analysis results, it can be concluded that the robotic arm was less understandable than the human arm while the fast-waved robotic arm was as understandable as the human arm. When it comes to the comfort ratings, the median value of the slow-waved robotic arm (2.0) was lower than that of the human arm (3.0) and the fast-waved robotic arm (3.0) in taking the road condition. In addition, the slow-waved robotic arm ($IQR_{\text{giving the road}}$: 2.0 – 4.0) was rated more scattered in comfort ratings than the human arm ($IQR_{\text{giving the road}}$: 3.0 – 4.0) and the fast-waved robotic arms ($IQR_{\text{giving the road}}$: 3.0 – 4.0), which indicated the higher uncertainty of participants in determining how comfortable the slow-waved robotic arm was. Combining with statistical analysis, we inferred that the slow-waved robotic arm was less comfortable than the human arm and the fast-waved robotic arm. It was noticeable that different types of arms showed no significant differences in the participants' politeness ratings. This further indicated that both robotic arms were perceived as polite as the human arm, regardless of how fast the arms waved. It might be because rather than taking actions directly without giving any signs, showing gestures was a polite behavior on the road no matter the gestures were transferred by which type of arms.

4.3 Limitations and future work

Several limitations should be considered when interpreting the results of this study. First, the present study about the effectiveness of robotic gestures on communication between the automated vehicle and other drivers was specific to some environments. For one thing, the premise of using the robotic arms to indicate gestures was the certain degree of visibility of the environment. When it comes to the physical environment with poor visibility (e.g., light conditions, angles) (Risto et al., 2017), the effectiveness of robotic arms may be limited. Under these circumstances, other features of robotic arms, such as luminance, should be considered. For another, the robotic arm might be needless when it comes to some simple traffic situations. For instance, turning on the automated vehicle's left indicator according to the usual traffic rule is sufficient for drivers behind the automated vehicle to understand its left-turn intention (Rodemerk et al., 2015). Therefore, in future studies, it is necessary to define the scenarios where the robotic arms would benefit in transferring the intentions of automated vehicles to adapt to traffic situations with various communication needs. Second, two typical gestures (i.e., "taking the road" and "giving the road") were primarily considered in this study. Other common gestures like left/right turns or gestures that expressed more complicated intentions of automated vehicles, such as thanks and warnings, can be included in future studies. In addition, the study adopted two general gestures that can keep consistently in different countries (Gupta et al., 2016). However, it should note that cultural differences exist in interpreting the intentions of some other gestures (Archer, 1997; Gupta et al., 2016; Stanciu et al., 2018). To make the gestures more universally applicable, future researchers should consider the localization or cultural differences in gestures when applying them to convey the intentions of automated vehicles. Third, the present study only considered the interaction between automated vehicles and human drivers in manual driving vehicles. Other road users, such as pedestrians and cyclists, are more vulnerable than drivers in cars and may require different communication patterns with automated vehicles (Anaya et al., 2014; Sewalkar & Seitz, 2019). Thus, the effect of gestures by robotic arms in communicating between automated vehicles and vulnerable road users is supposed to be investigated in future studies. Fourth, some practical issues related to the drivers in the automated vehicles should be concerned for the application of robotic arms. For one thing, the attitudes of drivers who own the automated vehicles to robotic arms (e.g., trust and acceptance) should be considered to determine how they would use this new device. For another, in situations with no communication needs, how to manage the presentation of robotic arms should be concerned because the existence of robotic arms may interfere with drivers' view in case of driving needs. In this circumstance, a self-adapted system, which can be embedded with the robotic arms systems to show up the robotic arm when it detects communication needs and put it down automatedly when there is no need for communications, might be a promising method. Fifth, this study adopted the simulated experiment based on video recordings for safety considerations. However, the external validity of the present study was limited, which may inhibit the generalization of the present findings. The communication between automated vehicles and humans in a true-to-life environment with the adoption of more realistic methods such as the Wizard of Oz for automation (Riek, 2012) can be considered in future studies.

5 Conclusion

This study proposed using robotic arms to transfer gestures on behalf of automated vehicles to communicate with human drivers in manual driving vehicles. We found that regardless of whether the gesture was giving the road or taking the road, transferring gestures by the fast-waved robotic arm exerted as fast and accurate responses as the human arm. In addition, the fast-robotic arm received as understandable, polite, and comfortable ratings from participants as the human arm. The results further indicated that conveying gestures by a fast-waved robotic arm (waving at 120 beats per minute) can facilitate effective communication between automated vehicles and human drivers in manual driving vehicles. The present study's findings provided implications for manufacturers and designers to adopt simple-device-based robotic arms to generate safe and user-friendly automated vehicles.

Acknowledgments

We appreciate the support from the National Natural Science Foundation of China with the Main Research Project on Machine Behavior and Human-Machine Collaborated Decision Making Methodology (Project Number: 72192820) and the Second Research Project on Human Behavior in Human-Machine Collaboration (Project Number: 72192824).

Disclosure of potential conflict of interest

The authors declare that they have no conflict of interest.

References

Alexander, R., Thompson, N., & Murray, D. (2017). Towards cultural translation of websites: a large-scale study of Australian, Chinese, and Saudi Arabian design preferences. *Behaviour & Information Technology*, *36*(4), 351-363.

Anaya, J. J., Merdrignac, P., Shagdar, O., Nashashibi, F., & Naranjo, J. E. (2014). Vehicle to pedestrian communications for protection of vulnerable road users. 2014 IEEE Intelligent Vehicles Symposium Proceedings,

Anderson, J. M., Nidhi, K., Stanley, K. D., Sorensen, P., Samaras, C., & Oluwatola, O. A. (2014). *Autonomous vehicle technology: A guide for policymakers*. Rand Corporation.

Aoki, S., Lin, C.-W., & Rajkumar, R. (2021). Human-robot cooperation for autonomous vehicles and human drivers: Challenges and solutions. *IEEE communications magazine*, *59*(8), 35-41.

Archer, D. (1997). Unspoken diversity: Cultural differences in gestures. *Qualitative sociology*, *20*(1), 79-105.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390-412.

Berger, J., Bacula, A., & Knight, H. (2021). Exploring Communicatory Gestures for Simple Multi-robot Systems. International Conference on Social Robotics,

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211-243.

Bradley, J. V. (1958). Complete counterbalancing of immediate sequential effects in a Latin square design. *Journal of the American Statistical Association*, *53*(282), 525-528.

Breslow, N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, *88*(421), 9-25.

Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement*, *56*(10), 15-23.

Clamann, M., Aubert, M., & Cummings, M. L. (2017). *Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles*.

Currano, R., Park, S. Y., Domingo, L., Garcia-Mancilla, J., Santana-Mancilla, P. C., Gonzalez, V. M., & Ju, W. (2018). ¡ Vamos! Observations of pedestrian interactions with driverless cars in Mexico. Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications,

De Clercq, K., Dietrich, A., Núñez Velasco, J. P., De Winter, J., & Happee, R. (2019). External human-machine interfaces on automated vehicles: effects on pedestrian crossing decisions. *Human factors*, *61*(8), 1353-1370.

Deo, N., Rangesh, A., & Trivedi, M. (2016). In-vehicle hand gesture recognition using hidden markov models. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC),

Deshmukh, A., Craenen, B., Foster, M. E., & Vinciarelli, A. (2018). The more I understand it, the less I like it: The relationship between understandability and Godspeed scores for robotic gestures. 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN),

Deshmukh, A., Craenen, B., Vinciarelli, A., & Foster, M. E. (2018). Shaping Robot Gestures to Shape Users' Perception: The Effect of Amplitude and Speed on Godspeed Ratings. Proceedings of the 6th International Conference on Human-Agent Interaction,

Do, W., Rouhani, O. M., & Miranda-Moreno, L. (2019). Simulation-based connected and automated vehicle models on highway sections: a literature review. *Journal of Advanced Transportation*, *2019*.

Epke, M. R., Kooijman, L., & De Winter, J. C. (2021). I see your gesture: a vr-based study of bidirectional communication between pedestrians and automated vehicles. *Journal of Advanced Transportation*, *2021*.

Faas, S. M., & Baumann, M. (2021). Pedestrian assessment: Is displaying automated driving mode in self-driving vehicles as relevant as emitting an engine sound in electric vehicles? *Applied ergonomics*, *94*, 103425.

Färber, B. (2016). Communication and communication problems between autonomous vehicles and human drivers. In *Autonomous driving* (pp. 125-144). Springer.

Fiedler, A., & Spelten, W. (2017). *Autonomous concept car smart vision EQ fortwo: Welcome to the future of car sharing*.

Flores, C., Merdrignac, P., de Charette, R., Navas, F., Milanés, V., & Nashashibi, F. (2018). A cooperative car-following/emergency braking system with prediction-based pedestrian avoidance capabilities. *IEEE Transactions on Intelligent Transportation Systems*, *20*(5), 1837-1846.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, *32*(200), 675-701.

Fuest, T., Schmidt, E., & Bengler, K. (2020). Comparison of methods to evaluate the influence of an automated vehicle's driving behavior on pedestrians: Wizard of Oz, virtual reality, and video. *Information*, *11*(6), 291.

Gleeson, B., MacLean, K., Haddadi, A., Croft, E., & Alcazar, J. (2013). Gestures for industry intuitive human-robot communication from human observation. 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI),

Guériau, M., & Dusparic, I. (2020). Quantifying the impact of connected and autonomous vehicles on traffic efficiency and safety in mixed traffic. 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC),

Gupta, S., Vasardani, M., & Winter, S. (2016). Conventionalized gestures for the interaction of people in traffic with autonomous vehicles. Proceedings of the 9th ACM SIGSPATIAL international workshop on computational transportation science,

Habibovic, A., Lundgren, V. M., Andersson, J., Klingegård, M., Lagström, T., Sirkka, A., Fagerlönn, J., Edgren, C., Fredriksson, R., & Krupenia, S. (2018). Communicating intent of automated vehicles to pedestrians. *Frontiers in psychology*, *9*, 1336.

Joisten, P., Liu, Z., Theobald, N., Webler, A., & Abendroth, B. (2021). Communication of Automated Vehicles and Pedestrian Groups: An Intercultural Study on Pedestrians' Street Crossing Decisions. In *Mensch und Computer 2021* (pp. 49-53).

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Citeseer.

Kantowitz, B. H., Roediger III, H. L., & Elmes, D. G. (2014). *Experimental psychology*. Cengage Learning.

Kesting, A., & Treiber, M. (2008). How reaction time, update time, and adaptation time influence the stability of traffic flow. *Computer-Aided Civil and Infrastructure Engineering*, *23*(2), 125-137.

Khan, S., & Tunçer, B. (2017). Intuitive and Effective Gestures for Conceptual Architectural Design: An Analysis Of User Elicited Hand Gestures For 3D CAD Modeling.

Lee, S., Son, Y., & Jin, J. (2010). Integrated human decision making and planning model under extended belief-desire-intention framework. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, *20*(4), 23.

Li, J., Currano, R., Sirkin, D., Goedicke, D., Tennent, H., Levine, A., Evers, V., & Ju, W. (2020). On-road and online studies to investigate beliefs and behaviors of Netherlands, US and Mexico pedestrians encountering hidden-driver vehicles. Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction,

Lohse, M., van Berkel, N., Van Dijk, E. M., Joosse, M. P., Karreman, D. E., & Evers, V. (2013). The influence of approach speed and functional noise on users' perception of a robot. 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems,

Mahadevan, K., Somanath, S., & Sharlin, E. (2018). Communicating awareness and intent in autonomous vehicle-pedestrian interaction. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems,

Mahmud, S. S., Ferreira, L., Hoque, M. S., & Tavassoli, A. (2017). Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs. *IATSS research*, *41*(4), 153-163.

Maurer, M., Gerdes, J. C., Lenz, B., & Winner, H. (2016). *Autonomous driving: technical, legal and social aspects*. Springer Nature.

May, A. D. (1990). *Traffic flow fundamentals*. Prentice-Hall.

Meier, A., Goto, K., & Wörmann, M. (2014). Thumbs Up to Gesture Controls? A Cross-Cultural Study on Spontaneous Gestures. International Conference on Cross-Cultural Design,

Merten, J. (1997). Facial-affective behavior, mutual gaze, and emotional experience in dyadic interactions. *Journal of Nonverbal Behavior*, *21*(3), 179-201.

Millard-Ball, A. (2018). Pedestrians, autonomous vehicles, and cities. *Journal of planning education and research*, *38*(1), 6-12.

Moon, A., Parker, C. A., Croft, E. A., & Van der Loos, H. M. (2013). Design and impact of hesitation gestures during human-robot resource conflicts. *Journal of Human-Robot Interaction*, *2*(3), 18-40.

Mukhtar, A., Xia, L., & Tang, T. B. (2015). Vehicle detection techniques for collision avoidance systems: A review. *IEEE Transactions on Intelligent Transportation Systems*, *16*(5), 2318-2338.

Oudshoorn, M., de Winter, J., Bazilinskyy, P., & Dodou, D. (2021). Bio-inspired intent communication for automated vehicles. *Transportation research part F: traffic psychology and behaviour*, *80*, 127-140.

Pheasant, S., & Haslegrave, C. M. (2018). *Bodyspace: Anthropometry, ergonomics and the design of work*. CRC press.

Pika, S., Nicoladis, E., & Marentette, P. (2009). How to order a beer: cultural differences in the use of conventional gestures for numbers. *Journal of Cross-Cultural Psychology*, *40*(1), 70-80.

Rahimian, P., O'Neal, E. E., Yon, J. P., Franzen, L., Jiang, Y., Plumert, J. M., & Kearney, J. K. (2016). Using a virtual environment to study the impact of sending traffic alerts to texting pedestrians. 2016 IEEE Virtual Reality (VR),

Rasouli, A., & Tsotsos, J. K. (2019). Autonomous vehicles that interact with pedestrians: A survey of theory and practice. *IEEE Transactions on Intelligent Transportation Systems*, *21*(3), 900-918.

Rettenmaier, M., Dinkel, S., & Bengler, K. (2021). Communication via motion–Suitability of automated vehicle movements to negotiate the right of way in road bottleneck scenarios. *Applied ergonomics*, *95*, 103438.

Riek, L. D. (2012). Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, *1*(1), 119-136.

Riek, L. D., Rabinowitch, T.-C., Bremner, P., Pipe, A. G., Fraser, M., & Robinson, P. (2010). Cooperative gestures: Effective signaling for humanoid robots. 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI),

Risto, M., Emmenegger, C., Vinkhuyzen, E., Cefkin, M., & Hollan, J. (2017). Human-vehicle interfaces: the power of vehicle movement gestures in human road user coordination.

Rodemerk, C., Winner, H., & Kastner, R. (2015). Predicting the driver's turn intentions at urban intersections using context-based indicators. 2015 IEEE Intelligent Vehicles Symposium (IV),

SAE International, (2021). Taxonomy and Definitions for Terms Related to Driving Automation Systems for on-Road Motor Vehicles. (J3016_202104). Retrieved from https://www.sae.org/standards/content/j3016_202104 (accessed April 30, 2021).

Saerbeck, M., & Bartneck, C. (2010). Perception of affect elicited by robot motion. 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI),

Salem, M., Kopp, S., Wachsmuth, I., Rohlfing, K., & Joublin, F. (2012). Generation and evaluation of communicative robot gesture. *International Journal of Social Robotics*, *4*(2), 201-217.

Sewalkar, P., & Seitz, J. (2019). Vehicle-to-pedestrian communication for vulnerable road users: Survey, design considerations, and challenges. *Sensors*, *19*(2), 358.

Sheikholeslami, S., Moon, A., & Croft, E. A. (2015). Exploring the effect of robot hand configurations in directional gestures for human-robot interaction. 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),

Sheikholeslami, S., Moon, A., & Croft, E. A. (2017). Cooperative gestures for industry: Exploring the efficacy of robot hand configurations in expression of instructional gestures for human–robot interaction. *The International Journal of Robotics Research*, *36*(5-7), 699-720.

Singer, T., Kobbert, J., Zandi, B., & Khanh, T. Q. (2020). Displaying the driving state of automated vehicles to other road users: An international, virtual reality-based study as a first step for the harmonized regulations of novel signaling devices. *IEEE Transactions on Intelligent Transportation Systems*.

Stanciu, S. C., Eby, D. W., Molnar, L. J., St. Louis, R. M., Zanier, N., & Kostyniuk, L. P. (2018). Pedestrians/bicyclists and autonomous vehicles: how will they communicate? *Transportation research record*, *2672*(22), 58-66.

Sucha, M., Dostal, D., & Risser, R. (2017). Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention*, *102*, 41-50.

Sullivan, J. M., Tsimhoni, O., & Bogard, S. (2008). Warning reliability and driver performance in naturalistic driving. *Human factors*, *50*(5), 845-852.

Wan, Y., & Sarter, N. (2022). Attention limitations in the detection and identification of alarms in close temporal proximity. *Human factors*, 00187208211063991.

Weber, F., Chadowitz, R., Schmidt, K., Messerschmidt, J., & Fuest, T. (2019). Crossing the Street Across the Globe: A Study on the Effects of eHMI on Pedestrians in the US, Germany and China. International Conference on Human-Computer Interaction,

Wilcoxon, F. (1992). Individual comparisons by ranking methods. In *Breakthroughs in statistics* (pp. 196-202). Springer.

Xing, Y., Lv, C., Cao, D., & Hang, P. (2021). Toward human-vehicle collaboration: Review and perspectives on human-centered collaborative automated driving. *Transportation research part C: emerging technologies*, *128*, 103199.

Xu, J., Broekens, J., Hindriks, K., & Neerincx, M. A. (2013). *Mood expression through parameterized functional behavior of robots*. IEEE.

Zandi, B., Singer, T., Kobbert, J., & Khanh, T. Q. (2020). International study on the importance of communication between automated vehicles and pedestrians. *Transportation research part F: traffic psychology and behaviour*, *74*, 52-66.

Zhang, J., Vinkhuyzen, E., & Cefkin, M. (2017). Evaluation of an autonomous vehicle external communication system concept: a survey study. International conference on applied human factors and ergonomics,

Zoghbi, S., Croft, E., Kulić, D., & Van der Loos, M. (2009). Evaluation of affective state estimations using an on-line reporting device during human-robot interactions. 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems,

About the Authors

Wei Zhang is a Ph.D. student at the Department of Industrial Engineering, Tsinghua University. Her research interests include human-computer interaction in automated driving, system design, and human performance improvement.

Changxu Wu is a full professor at the Department of Industrial Engineering, Tsinghua University. His research interests include integrating cognitive science and engineering system design, especially modeling human cognition systems and human behavior with its applications in system design, improving transportation safety, and promoting human performance and safety in human-machine interaction.

Xianwen You is a Ph.D. student at the Department of Industrial Engineering, Tsinghua University. His research interests include human performance modeling and human-computer interaction.

Kust Leo received his master's degree in industrial engineering at the State University of New York at Buffalo. His research interests include human-vehicle interactions and system design.

Yijie Chen is a Ph.D. student at Arizona State University. His research interests include human-vehicle interactions and system design.

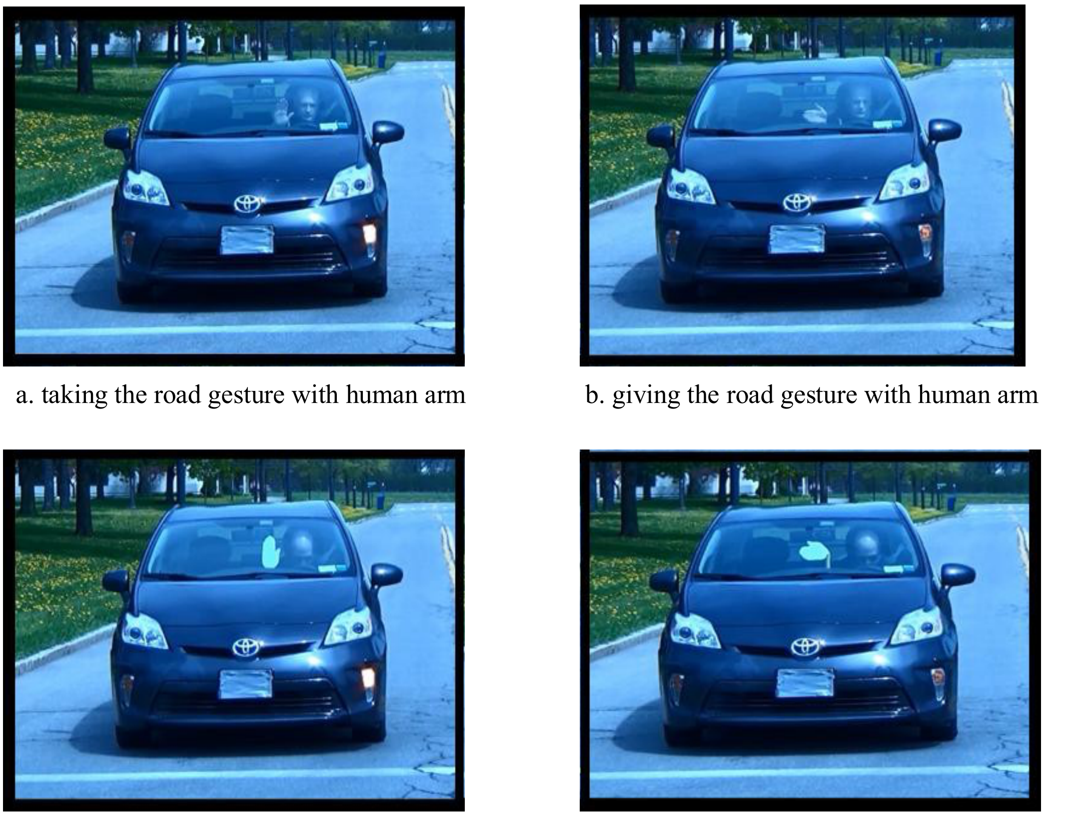
Linlei Shi is a Ph.D. student at the College of Computer Science and Technology, Zhejiang University. His research interests include human-computer interaction, user experience, augmented reality, and vibrotactile warnings.

Appendix

Table A1. balanced Latin Square for six experimental conditions

Order of testing conditions						
subjects	**1 st**	**2 nd**	**3 rd**	**4 th**	**5 th**	**6 th**
a	1	2	6	3	5	4
b	2	3	1	4	6	5
c	3	4	2	5	1	6
d	4	5	3	6	2	1
e	5	6	4	1	3	2
f	6	1	5	2	4	3
(repeat)						

Note: 1 - video 1: taking the road gesture by the human arm; 2 - video 2: giving the road by the human arm; 3 - video 3: taking the road by the fast-waved robotic arm; 4 - video 4: giving the road by the fast-waved robotic arm; 5 - video 5: taking the road by the slow-waved robotic arm; 6 - video 6: giving the road by the slow-waved robotic arm.



a. taking the road gesture with human arm

b. giving the road gesture with human arm

c. taking the road gesture with the robotic arm

d. giving the road gesture with the robotic arm

Figure 1. Gesture transferring scenarios. a. "taking the road" gesture with human arm; b. "giving the road" gesture with human arm; c. "taking the road" gesture with the robotic arm; d. "giving the road" gesture with the robotic arm (the speed of the gestures transferred by robotic arms cannot be presented in photos)

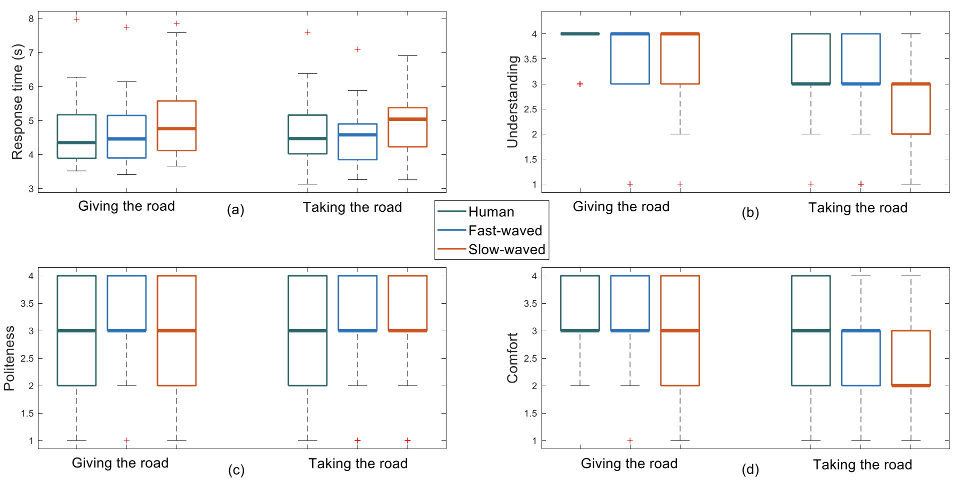


Figure 2. Boxplots of raw data of dependent variables in terms of arm type and gesture type, (a) response time, (b) understanding (1 = "not understandable at all", 4 ="very understandable"), (c) politeness (1 = "not polite at all", 4 = "very polite"), (d) comfort (1 = "not comfort at all", 4 = "very comfort") (the thick middle line within the box represents the median value, the lower and upper hinge of the box represent the first (Q1) and third quartile (Q3), whiskers were no larger than 1.5 times the interquartile range, and red plus signs represent outliers).

Table 1. Descriptive statistical data of response time and response accuracy for different arms and gestures

Gesture type	Arm type	Response time (s)	Response accuracy (%)
Giving the road	human	4.476 (0.783)	100 (0.00)
	Fast-waved robotic	4.519 (0.762)	93.3 (25.4)
	Slow-waved robotic	4.874 (0.977)	90.0 (30.5)
Taking the road	human	4.493 (0.871)	96.7 (18.3)
	Fast-waved robotic	4.386 (0.693)	96.7 (18.3)
	Slow-waved robotic	4.867 (0.931)	86.7 (34.6)

Notes: Data were mean (standard deviation).

Table 2. Median, interquartile range, and mean rank of subjective ratings for different arms and gestures.

Gesture type	Arm type	Understanding			Politeness			Comfort		
		Median	IQR	Mean rank	Median	IQR	Mean rank	Median	IQR	Mean rank
Giving the road	human	4.0	4.0-4.0	2.2	3.0	2.0-4.0	2.02	3.0	3.0-4.0	2.22
	Fast-waved robotic	4.0	3.0-4.0	1.93	3.0	3.0-4.0	2.03	3.0	3.0-4.0	2.05
	Slow-waved robotic	4.0	3.0-4.0	1.87	3.0	2.0-4.0	1.95	3.0	2.0-4.0	1.73
Taking the road	human	3.0	3.0-4.0	2.15	3.0	2.0-4.0	1.93	3.0	2.0-4.0	2.43
	Fast-waved robotic	3.0	3.0-4.0	2.22	3.0	3.0-4.0	2.07	3.0	2.0-3.25	2.15
	Slow-waved robotic	3.0	2.0-3.25	1.63	3.0	3.0-4.0	2.00	2.0	2.0-3.0	1.42

Note: IQR: interquartile range; for understanding: 1 = “not understandable at all”, 4 = “very understandable”); for politeness: 1 = “not polite at all”, 4 = “very polite”; for comfort: 1 = “not comfort at all”, 4 = “very comfort”.

Table 3. The Wilcoxon test results of understanding and comfort ratings for different arms under different gestures.

Gesture type	Arm type comparison	Understanding			Comfort		
		Z	*p*	*r*	*Z*	*p*	*r*
Giving the road	Fast-waved vs. Human	-1.930	0.054	-0.35	-1.291	0.197	-0.24
	Slow-waved vs. Human	-2.310	0.021*	-0.42	-2.500	0.012*	-0.46
	Slow-waved vs. Fast-waved	-0.272	0.785	-0.05	-2.066	0.039*	-0.38
Taking the road	Fast-waved vs. Human	-0.302	0.763	-0.06	-1.698	0.090	-0.31
	Slow-waved vs. Human	-2.556	0.011*	-0.47	-4.234	<.001***	-0.77
	Slow-waved vs. Fast-waved	-2.982	0.003*	-0.54	-3.578	<.001***	-0.65

Note. *p < 0.05; ***p < 0.001.