Pedestrians' Understanding of a Fully Autonomous Vehicle's Intent to Stop: Utilizing Video-based Crossing Scenarios

Michal Hochman[1], Yisrael Parmet[1], Tal Oron-Gilad[1]

Industrial engineering and Management, Ben-Gurion University of the Negev, Israel, hochmanm@post.bgu.ac.il

Abstract

Background. External human-machine interfaces (eHMI) indicate Fully Autonomous Vehicles' (FAVs) intents, contributing to their communication with pedestrians. We still do not know enough about how eHMI propositions lead pedestrians to comply in conflicting situations.

Objective. Findings on fixed crossing scenes suggest that pedestrians' decision-making depends on the eHMI implementation and the 'vehicle's distance from the crossing. We aim to enhance this work, looking at dynamic crossing situations.

Method. Thirty-four adult participants observed 56 road-crossing video scenarios as if they were pedestrians intending to cross. A single FAV drove at 40 km/h. Scenarios differed by car size, eHMI message type, and the FAV's initial distance from the crossing place. Participants had to decide whether to cross or not by pressing designated buttons. Following each scenario, their subjective Understanding of the FAV's intention was obtained. Decision measurements and eye-tracking data were collected.

Results. Eye-tracking data confirmed that all pedestrians fixated on the eHMI, yet only 53% of the responses were compatible with its proposition. More incompatible responses were observed for the close distance. An interaction between distance and eHMI proposition revealed that when the eHMI indicated participants to cross, and the FAV's initial location was close, most participants decided not to cross. Distance influenced participants' response time; pedestrians decided faster in the closer distance. Overall, subjective Understanding of the FAV's intention was low.

Conclusion. Using video-based scenarios, we showed the combined effect of context and eHMI meaning on pedestrians' crossing decisions. Relative to fixed scenes, pedestrians were more conservative and relied less on the eHMI suggestions. Interactions of distance and message meaning affected compatibility and response time. Even when pedestrians understood the eHMI message, they did not necessarily comply. Distance of the vehicle from the crossing place influenced the crossing decision, as it does today.

Keywords: Pedestrian behavior, external Human-Machine Interfaces (eHMI), distance from the crossing place

1. Introduction

External human-machine interfaces (eHMI) can indicate Fully Autonomous Vehicles' (FAVs) intent to stop. They contribute to the communication with pedestrians by reducing the uncertainty regarding FAV intents and improving pedestrians' initial trust and Understanding (Clamann, Aubert, and Cummings, 2017; Ackermans et al., 2020; Kaleefathullah et al., 2020; Deb, Strawderman, and Carruth, 2018). Research revealed that pedestrians tend to look at the eHMI before making the crossing decision; however, when they comply with the eHMI proposition, they do not necessarily decide according to its features. Instead, they tend to hesitate before making the decision (Hochman et al., 2020). Pedestrians' decisions to cross depend on the eHMI and the crossing conditions, e.g., vehicles' distance from the crossing place (Robinette, Wagner, and Howard, 2013; Hochman et al., 2020; Mahadevan, Somanath, and Sharlin 2018; Tabone et al., 2021).

Understanding of the eHMI can be measured in several ways. One measure is the compatible rate, whether the pedestrian's decision was in agreement with what was displayed on the eHMI (Ackermann et al., 2019). Another way is through subjective questionnaires (Deb, Carruth, and Strawderman, 2020; Tabone et al., 2021). The third way is via the decision-making time. Previously, in fixed scenes, we found that decision-making was fastest when the eHMI suggestion was to stop, especially at a close distance (Hochman et al., 2020). When the eHMI proposition conflicted with pedestrians' expectations, they gazed at the eHMI to understand the message and took more time to decide (Hochman et al., 2020).

1 Corresponding author: Michal Hochman. Tel.: ++972-8-6461434

E-mail address: hochmanm@post.bgu.ac.il

Additionally, a negative correlation was reported between pedestrians' subjective Understanding of the FAV intention and gaze fixation duration (Liu et al., 2020). The stopping profile, i.e., the FAVs' braking pattern, was also reported to affect pedestrian behavior and was suggested as a factor affecting pedestrians' head-turning patterns (Lv et al., 2021). Closer, more immediate FAV stopping distance from the pedestrian's crossing place was less confusing for the pedestrian and led to less head-turning. However, it caused the most extended response times (Lv et al., 2021). All in all, we know less about whether pedestrians understand the eHMI messages in various crossing conditions and how they affect their crossing decisions.

The current study examines the robustness of previous findings on fixed crossing scenes (Hochman et al., 2020) for dynamic crossing situations. Noting that pedestrians' decision-making depended on the eHMI implementation and the car distance, we now examine factors related to the crossing context; distance from the crossing place, car size, stopping profile, and the eHMI contextual- message meaning and message type. The following hypotheses are suggested: H1- following Deb et al. (2018), we expected that response time would increase in the close distance when the FAV suggestion is to cross. We foresee the opposite in the far distance. H2- following Hochman et al. (2020), in case of a conflict with the eHMI suggestion, pedestrians will respond slower or against the 'FAV's suggestion, even though they understand its intention. H3- We expect that the stopping profile will affect the decision to cross. Pedestrians will notice the difference in stopping profile, but we are not sure how it will affect their decision.

2. Methodology

2.1. Participants

Thirty-four participants (22-31 years old) observed 56 urban road crossing scenarios as if they were pedestrians intending to cross (Figure 1). Data of one participant was excluded due to technical problems. As compensation, 24 participants received course credit, and 10 received a payment of ~$10. All participants had normal contrast sensitivity and visual acuity of at least 6/6.

2.2. Apparatus

2.2.1. Experimental Environment

The study was conducted using a dedicated apparatus and a 42" screen at the Pedestrian behavior lab in the Industrial Engineering and Management Department at the Ben-Gurion University of the Negev. Participants were standing in front of the screen, two-meter away, and wearing the head-mounted eye tracker (see Figure 1).

Figure 1: The experimental testbed, consisting of a 42" screen showing the crossing scenario video, the Dikablis eye-tracking system, and response buttons to collect crossing decisions.

2.2.2. Eye-tracking system

A head-mounted eye tracker was used to measure pupil diameter and gaze direction with an accuracy visual angle of 0.5 degrees and sample rate of 60Hz (Dikablis Glasses, Ergoneers GmbH, Geretsreid, Germany).

2.2.3. Video-based scenarios generation

We generated fifty-six crossing scenario videos using the VT-MAK VR tools (https://www.mak.com) with a 3D terrain model of a typical local city. The crossed road was a one-directional one-lane urban road. To add realism to the scene, the typography of the city included buildings, light posts, vegetation, etc. (Figure 2). Videos were filmed from the pedestrian's perspective as if standing on the curb and looking to the left or right before crossing the street. Each scenario included a combination of a single FAV (small or big) driving at 40 km/h, and the FAV's initial distance from the crossing area (far – 7 sec or close – 3 sec) (Figure 3). The eHMI was located on the car's roof, which was advantageous in previous research (Bazilinskyy & Dodou, 2020). It included a sign conveying either a status message ('Slowing' or 'Driving') or an advice message ('Cross' or 'Stop') in Hebrew. When the FAV intended to stop, it started slowing and stopped either closer (~ 1 meter) or more abruptly further away from the pedestrian(~ 2 meters) (totaling two stopping profiles). Besides, baseline videos without the eHMI were created, with a variation of car size and crossing distance (see Appendix A for the entire video content). The eHMI initial size was 1*1 cm and a 6.36-degree horizontal visual angle in close distance. The initial size was 2.1*2.1cm in the far distance with a 6.01-degree horizontal visual angle.

2.3. Road Crossing Task

Each participant took part in eight consecutive short sessions. In each session, they observed seven crossing video-based scenarios (totaling 56 scenarios). Participants had to decide as quickly as possible whether they would cross the road or not. The decision was registered by pressing the designated green button - 'I will cross' or the red button- 'I will not cross'. Following each scenario, the participant answered four questions regarding the clarity of the FAV's intention (see subjective questionnaire).

A

B

C

D

Figure 2. Sample crossing scenes, as seen from the perspective of the pedestrian. Each row (a-d) demonstrates an examined factor. a. Message type: left- Status, right- Advice (in Hebrew Slowing/Cross!) b. distance: left- close, right- far c. car size: left - Small (Kancil), right- Large (Audi), d. Stopping profile: left-longer more moderate stopping, right- close more immediate stopping.

Figure 3: Sample crossing scenes, as seen from the pedestrian's perspective. The initial distance of the FAV at the beginning of the scenario is demonstrated: left- close (~3 seconds to the crossing place), right- far (~7 seconds).

2.4. Dependent Variables

2.4.1. Compatibility of the participant's decision with the eHMI sign meaning

The participant's decision (to cross or not) with the eHMI sign meaning was defined as a binary variable - incompatible (or error) if the participant's decision did not correspond to the sign's proposition or compatible if it corresponded. In the statistical model, we predicted the estimated error probability for compatibility. Videos with no eHMI were excluded as error probability cannot be defined.

2.4.2. Response Time

RT is defined as the time from the moment the video was initiated until the participant pressed a decision button. RT was analyzed only for the compatible responses since it was interesting to compare the time it took to respond to the crossing complexities.

2.4.3. Subjective Measurements

2.4.3.1. The intermediate questionnaire

Following each scenario, participants answered four questions regarding the clarity of the FAV's intention. The intermediate questionnaire: 1. Did you understand the FAV message? (yes/no). 2. How confident are you in your decision? (On 10-point rating scale). 3. To what extent did you experience the situation as dangerous? (On 10-point rating scale). 4. What was the FAV's intention (short free text answer).

Once the experimental scenarios ended, we coded participants' words and responses to question 4 (such as: *tell me to stop/ cross, intent on driving/stop, keep driving*) into their two meanings (1= don't cross, 2= cross) and analyzed whether there was a difference between the subjective answers and the FAVs suggestion (we termed this "subjective compatibility"). For example, suppose a participant stated that the car intended to stop, but FAVs suggestion was not to cross; in this case, the participant's "subjective compatibility" was 0, and if the FAV's intention meant to stop, the subjective compatibility was 1. In addition, we compared pedestrians' subjective Understanding with their actual decision. For example, if a participant said that she thought that the car suggested crossing, the eHMI suggestion was to cross, and her actual response was to cross. Her Understanding and her response both fit the meaning of the eHMI.

2.4.3.2. Final interview

At the end of each experiment, an open interview was conducted to learn more about how each participant made their crossing decisions.

2.4.3.3. Final questionnaires

After the interview, two final questionnaires were administered. The Sensation Seeking Scale and (SSS -V) (Zuckerman, 2007), and the Technology Assessment Propensity (TAP; Ratchford & Barnhart, 2012)

2.4.4. Eye-tracking measurements

Eye movements and fixations data were collected and synced with the experimental timeline for each crossing scenario. Once the experiment ended, we determined manually, for each scenario, whether there was at least one fixation within the defined area of interest (AOI), that is, on the eHMI or not. A fixation was defined as a period of at least 100 msec that the eyes remain relatively still. If there was at least one fixation, it was registered as if the participant looked at the eHMI before making the crossing decision.

2.4.5. Experimental Design

The study followed a within-subject repeated measures design. The effect of six independent variables eHMI (included/none-control), message type (status message/advise message), message suggestion (cross / don't cross) stopping profile (none/ close/far), car size (big/ small) and car distance (close/far) were evaluated by four crossing decision related measures.

2.5. Procedure

Participants were invited individually to the lab for a ~45 minutes meeting. First, they performed visual acuity and contrast sensitivity tests (Ginsburg, 1984). Next, the eye calibration was done. After calibration, participants completed a short road-crossing task with three video scenarios with no eHMI. Following this, the experiment was divided into eight consecutive short sessions. After each session, there was a short break. The sessions and the videos within them were given in random order. Sessions included videos with all combinations of car size, distance from the crossing place, and eHMI message type. Each session contained one baseline video.

Throughout the experiment, each video variation appeared two times with slight variations of the surrounding urban crossing road environment (e.g., building facade). After each video, the participant answered questions regarding the clarity of the FAV's intention. At the end of the experiment, an open interview was conducted.

2.6. Data Analysis

The two dependent variables, ***error probability*** (responses incompatible with the eHMI meaning) and ***response ti me***, were analyzed within the Generalized Linear Mixed Model (GLMM) framework. The error probability varia ble is binary. Therefore, logistic regression was used; a standard linear model analyzed response time (after log-t ransformation to achieve normality). The initial fixed factors included in the model were message type, stop-ping profile, car size, distance, and interactions in both models. The random-effects factors were the participants and t he scenarios to account for individual differences among participants and scenarios. In both cases, the outcomes presented in the manuscript are of the final model after implementing the backward elimination algorithm. In the backward elimination process, the main effects involved in the interaction were not considered for elimination. P ost hoc analyses (Holm–Bonferroni) were performed on the factors that came out significant and remained in the final model.

3. Analysis and Results

Using the eye tracker data, we validated whether a participant gazed on the eHMI at least once in each scenario. We confirmed that pedestrians fixated on the eHMI at least once in all cases (100%). Yet, only 53% (861 out of 1613) of the responses were compatible with the eHMI's proposition. The number of incompatible responses (752) was significantly higher for the close distance 59% (447) than the far 41% (305). Also, there was an interaction between distance and the eHMI intent. When the eHMI proposed participants to cross, and the FAV's initial location was close, only 19% of the responses were compatible; most pedestrians decided not to cross despite the eHMI's suggestion. The opposite occurred in the far distance, where 76% of responses were compatible when participants were encouraged to cross.
Generally, for all scenarios, it was found that pedestrians made their decision (to cross or not) before the FAV completely stopped when the suggestion was to cross (regardless of initial distance). Therefore, they did not observe the point in the scenario where the two stopping profiles occurred. Hence, the main difference between the stopping profiles was when the 'FAV's suggestion was to stop, and the car did not stop at any point.

3.1. Error probability (incompatibility with the eHMI proposition)

The final GLMM for error probability is given in Table 1.

Table 1. The effect of eHMI related factors (message meaning, message type) and crossing context factors (distance, stopping profile) on the error probability as derived from the GLMM.

Factors	F (1, 865)
Distance	40.86***
Stopping profile	2.77***
Message type	4.09
Stopping profile* Message type	1.27*
Stopping profile*Distance	119.30***

The F ratio measures the overall significance of the differences between the parameters within the model (distances, stopping profiles, two message types). *p<.05, **p<.0.01, ***p<0.001. The stopping profile in the model refers to two stopping profiles for the proposition to cross and a one-stop proposition for when the FAV did not stop.

3.1.1. Distance and Stopping profile

Post hoc analyses revealed that in the close distance, the error probability (incompatible responses) was higher for both stopping profiles (far stopping profile (estimated probability = 0.87, SE = 0.03) and close stopping profile (estimated probability = 0.85, SE = 0.04) when the message meaning was to cross, compared to when the message suggested not to cross, p<0.0001. In the close distance, when the message suggested not to cross, there were almost zero error (estimated probability = 0.03, SE = 0.01) (Figure 4, right). In the far distance the 'don't cross' suggestion (that is - 'driving' or 'stop') (estimated probability = 0.66, SE = 0.06) error probability was significantly higher compared to both stopping profiles stopping (far stopping profile (estimated probability = 0.20, SE = 0.04), close stopping profile (estimated probability = 0.18, SE = 0.04)), p<0.0001.

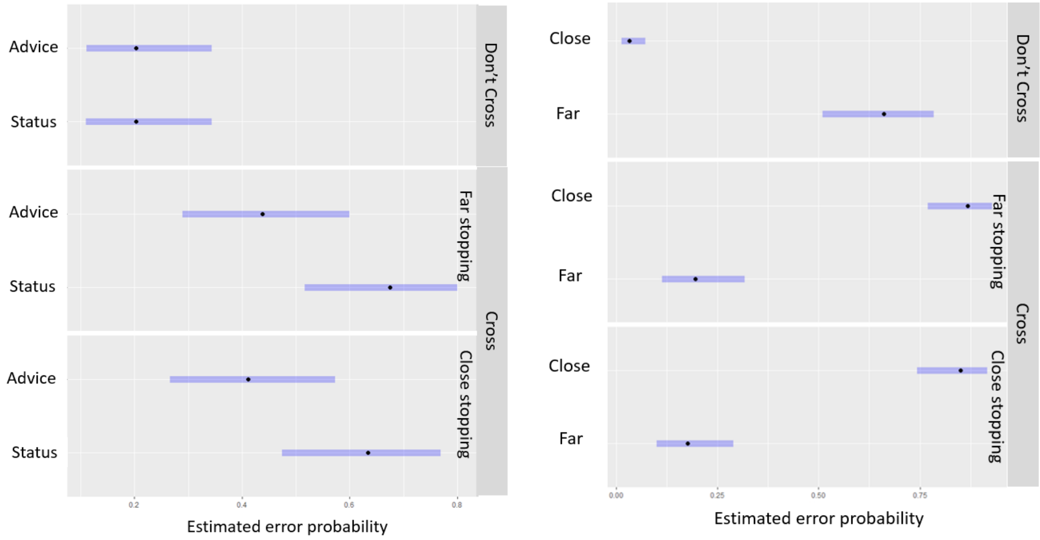

Figure 4: Error probability: Left-by message type and stopping profile, Right – by distance and stopping profile.

3.1.2. Message type and stopping profile

Post hoc (Holm method) analyses revealed a significant difference between eHMI suggestions in both stopping profiles and "don't cross" suggestion (Figure 4, left), p<0.01. In cases the suggestion was "don't cross" the error probability was very low and equal for both message types (advice (estimated probability = 0.2, SE = 0.05), status (estimated probability = 0.2, SE = 0.05)) and significantly lower than both stopping profiles in both message types (far stopping profile: advice (estimated probability = 0.44, SE = 0.07, p<0.01), status (estimated probability = 0.67, SE = 0.06, p<0.0001), (close stopping profile: advice (estimated probability = 0.41, SE = 0.07, p<0.01), status (estimated probability = 0.63, SE = 0.06, p<0.0001)). In addition, the estimated error probability was significantly higher in the status message compared to the advice message for both stopping profiles, P<0.001. Also, there was a significant difference between both message type in both stopping profile, p<0.01.

3.2. Response time

The significant effects included in the final model for response time were the eHMI related factors, message type, and crossing context factors (distance and stopping profile), see Table 2.

Table 2. The effect of eHMI related factors (message type) and crossing context factors (distance, stopping profile) on the estimated response time (ln transformed) for the compatible responses (GLMM)

Factors	F (1, 865)
Distance	190.08***
Stopping profile	16.28***
Message type	0.11
Stopping profile* Distance	11.97 ***
Message type*Distance	6.50*

The F ratio measures the overall significance of the differences between the parameters within the model (distances, stopping profiles, two message types). *p<.05, **p<.0.01, ***p<0.001.

3.2.1. Distance and stopping profile

Results revealed that distance influenced participants' reaction time for the compatible responses. In the closer distance (RT= 1.90 sec, SE=1.07), pedestrians tended to decide quickly than in the far distance (RT= 2.77 sec, SE=1.07), p<0.001 (Figure 5, right). Also, a significant interaction was found between the stopping profile and distance, p<0.001. Post hoc (Holm method) analyses revealed that in the close distance response time, there were no significant differences between both stopping profiles (far stopping profile (RT= 1.90 sec, SE=1.08) and close stopping profile (RT= 1.88 sec, SE=1.08) and when message suggested not to cross (RT= 1.92 sec, SE=1.07), p>0.05, (Figure 5, right). However, in the far distance, it was found that profile response times were almost equal (far stopping profile (RT= 2.48 sec, SE=1.07), close stopping profile (RT= 2.56 sec, SE=1.07), significantly lower than when the eHMI suggestion was not to cross (RT= 3.35 sec, SE=1.01), p<0.001. These findings imply that in the far distance, pedestrians tend to cross; therefore, when the suggestion was not to cross, they hesitated before implying the eHMI suggestion, compared to when the suggestion was to cross. However, pedestrians tend not to cross in close distance, and when suggested to cross, they hesitate before implying the suggestion. In addition, results revealed no distinguishing difference between the two stopping profiles.

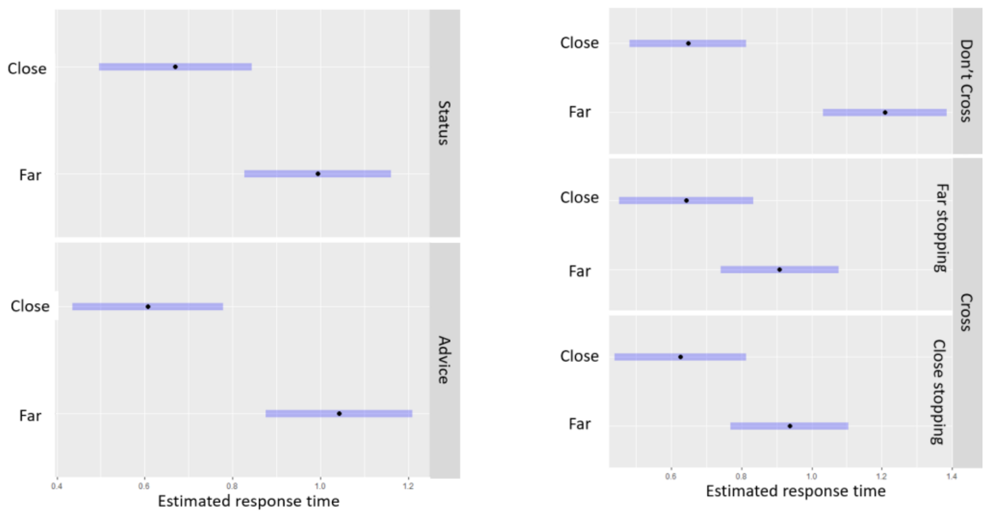

Figure 5: Response time for compatible responses: Left- by distance and message type, Right - by distance and stopping profile

3.2.2. Message type and distance

There was a significant interaction in the response time between the message type and FAVs distance (p<0.05) for compatible responses, p<0.0001(Table 2, Figure 5, left). Overall, response times were higher for both message types in the close distance than in the far distance (p<0.0001). Post hoc (Holm method) analyses revealed that the response time for the status message was higher in the far distance (RT = 2.7 sec, SE = 1.07) compared to the close one (RT = 1.95 sec, SE = 1.08), (p<0.0001). In addition, for the advice message, the far distance (RT = 2.83 sec, SE = 1.07) was significantly higher compared to the close distance (RT = 1.82, SE = 1.07), (p<0.0001).

3.3. Pedestrian's subjective Understanding

Recall it was found that there were compatible responses in 53% of the cases; however, even within those, only 33% fit the perceived message meaning, subjective Understanding, and the actual pedestrian response (Table 3). The other compatible responses (20%) were not based on a complete understanding of FAV intention. In these cases, the pedestrians decided to respond compatible due to a lack of Understanding (subjective understanding=0 that occurred in 13% of all responses) or when they noted that they did not understand the meaning or had a conflict or misunderstanding (in 7% of all responses), mainly in cases where FAV was far away. In addition, there were incompatible responses (19% of all cases) that occurred with correct Understanding, which means the pedestrian knew the FAV's intention but decided to respond against its suggestion. Moreover, in cases of Understanding, incompatibility occurred mainly in scenarios when the FAV was close and the suggestion was to cross (78%) or when FAV was far and the suggestion was not to cross (92%). A similar trend occurred when the pedestrian did not understand the FAV. They mainly decided not to cross in the close distance when suggested to cross (77%) and cross in the far distance when suggested not cross (95%). Also, when pedestrians noted that they did not understand the meaning or had a conflict, they had more incompatible responses in the close distance (18% - when summing both cases*) than the far distance (10%). In the compatible responses, one can see an opposite trend; in the far distance, there were more compatible responses (13%) compared to the close distance (6%).

Table 3. Breakdown of responses per pedestrians' Understanding of the message meaning and distance

Pedestrians' Understanding of message meaning	**Compatible responses** **All** **By initial distance** **far and close**		**Incompatible responses** **All** **By initial distance** **far and close**		**Total**
Understanding fits message meaning	33% (537)		19% (306)		**52%**
	Far 15%	Close 18%	Far 9%	Close 10%	
Understanding **does not** fit message meaning	13% (203)		17% (275)		**30%**
	Far 9%	Close 4%	Far 6%	Close 11%	
Did not understand the meaning / Had a conflict with FAV suggestion and its' behavior*	7% (121)		11% (171)		**18%**
	5%	2%	4%	7%	
Total	**53% (861)**		**47% (752)**		**100%**

* For example, the FAV suggestion was to cross, but it accelerated.

4. Discussion

Following a previous fixed scenario study (Hochman et al. 2020), we aimed to continue focusing on the interaction between crossing decisions and the eHMI suggestion in dynamic crossing scenarios. We explored which parameters affect pedestrian Understanding of the FAV's intentions in two crossing situations while varying eHMI meanings. Relative to (Hochman et al. 2020), the eHMI displays were simplified by removing color conversions and the symbol modality and leaving only neutral cyan text messages. This allowed us to focus on the dynamic properties of the crossing decision and the sense of *urgency to decide* that the dynamic settings create and examine FAV stopping profiles. Overall, pedestrians were more conservative in the dynamic context and relied less on the eHMI suggestions than what was reported for fixed static crossing scenes previously, 53% versus 75% compatible responses. However, there were twice as many crossing suggestions (to cross) in this research due to the stopping profiles manipulations than not crossing ones (do not cross), which may have impacted participants' responses.

Looking at the incompatibility responses and perceived Understanding of the eHMI proposition, there are two unwanted use cases; 1) when the pedestrian does not understand the FAV's intention in a specific crossing situation (misunderstanding), and 2) when the pedestrian understands the FAV's intention in context but acts against it, implying on difficulties in trusting the FAV's suggestion (mistrust of the FAV). Both mistrust and misunderstanding caused incompatibility, mainly in scenarios where the FAV was close and the suggestion was to cross or when FAV was far and the suggestion was not to cross. Still, these two situations are not equal, as the far distance was about 7 seconds away from the pedestrian, and it is considered safe to cross at such a gap. Nevertheless, if pedestrians hesitate to cross when the FAV is close to them, mistrust can affect traffic flow, as all FAVs will have to reach a complete stop before the pedestrian aims to cross.

4.1. Error probability

An interaction between distance and message meaning was found when analyzing error probability. Overall, in the far distance, it was found that when eHMI suggestion was to cross ('Slowing' or 'Cross'), the error probability for compatible responses was significantly lower than when the eHMI suggestion was not to cross ('Driving' or 'Stop'), $p<0.001$, (Table 1, Figures 4, right & 5, right). In the close distance, the error probability was significantly higher when the suggestion was to cross than when the message meaning was not to cross, where almost all responses were compatible (Table 1 & Figure 4, right). These results imply that in close distance, pedestrians tend not to cross. Therefore, there were almost no incompatible responses when the suggestion was to cross; however, in the far distance, the pedestrian tends to cross, so there were more incompatible responses when suggested not to cross. These findings confirm h1 and align with previous research (Hochman et al., 2020) that explored the effect of distance on pedestrians' crossing decisions (Clamann, Aubert, and M. L. Cummings, 2017; Winter and Shyrokau, 2019). When suggested not to cross in the far distance, they sometimes take a risk and cross against the eHMI proposition and therefore have more incompatible responses. In the close distance, pedestrians seemed to be more conservative when they were suggested crossing, and therefore they did not always risk crossing against their expectations. These findings align with previous research (Hochman et al. 2020) and can be explained by pedestrians' conservative trend not to cross. In addition, the estimated error probability was significantly higher in the status message than the advice message for both stopping profiles, $P<0.001$which means that the advice message was probably more understandable than the status message.

4.2. Response time

Results revealed that distance influenced participants' reaction time for the compatible responses. In the closer distance, pedestrians tended to decide quickly than in the far distance (Table 2). Analyses revealed no significant differences between both stopping profiles in the close distance (Figure 5, right). These findings are in line with the fixed scene research (Hochman et al., 2020). However, in the far distance, it was found that both profile response times were almost equal but significantly lower than when the eHMI suggestion was not to cross, $p<0.001$. This finding partly confirms h1; that is, response time would increase in the far distance when the FAV suggestion was not to cross compared to when the suggestion is to cross. However, there was no significant difference in the response time between the two message meanings in the close distance, unlike our expectations. These findings imply that in the far distance, pedestrian tends to cross; therefore, when the suggestion was not to cross, the pedestrian hesitated before implying the eHMI suggestion compared to when the suggestion was to cross. Also, results revealed no distinguishing difference between the two stopping profiles regarding response time or error probability. This finding contradicts our expectations and previous research (Lv et al., 2021). This might be due to the profiles not being distinguished enough. In addition, results revealed a significant difference in the response time for the status message between the far and close distance. Also, there was a significant difference between the far and close distance ($p<0.001$). In both cases, the response time for the far distance was higher than the close distance. These results can be explained by the fact that the pedestrian needed to wait before deciding in the far distance since he waited for the FAV to get closer to understand its intention and eHMI.

4.3. Subjective Understanding

Although it was found that there were 53% compatible responses, only 33% fitted the message meaning, subjective Understanding, and actual pedestrian response (Table 3). In the other compatible answers (that were 13% of all answers), the pedestrians decided to respond compatible due to lack of Understanding (subjective Understanding was 0, that is, the Understanding did not fit the meaning), or when noted they did not understand the meaning or had a conflict (totaling 7% of all answer), mainly in cases where FAV was far. In these cases, the reason for decision making can be other factors like distance affected the crossing decisions, as it does today (Hochman et al. 2020; Tabone et al. 2021). Overall, when they did not understand the meaning or had a conflict, they had more incompatible responses in the close distance than the far distance. It implies that they felt less safe to take a risk at a close distance when they did not understand FAV intention. These findings confirm h2, and in line with previous research (Hochman et al. 2020), when crossing environments conflict with eHMI suggestion, the pedestrian will respond slower or against the FAV suggestion, though he understood the FAV intention. We saw that sometimes Understanding is not always sufficient to make the compatible decision, and sometimes in cases of lack of Understanding, pedestrians choose the compatible response. In both cases, the road complexity might affect the decision-making more than understanding the FAV.

After analyzing the results, we can say that even though the eHMI displays were simplified, there are still cases of misunderstanding or conflicts in specific scenarios where the pedestrian decided to cross despite the eHMI suggestion. These decision-makings occurred appeared mainly in the far distance.

5. Conclusions, limitations, and future directions

Looking at the two unwanted use cases of misunderstanding and mistrust and focusing on the one that occurred when the FAV was close, we argue that if such mistrust and misunderstanding were observed when conducting a lab experiment that contained only one FAV and one pedestrian, a higher and more significant trend could be expected in the real world. Therefore, it is essential to provide e-HMI designs and encourage the FAV behavior to minimize error probability by increasing pedestrians' trust in the FAV. They currently mainly consider the 'FAV's distance from the crossing place in conflict situations, as they do today. Moreover, we suggest adjusting the FAV behavior in those situations to increase pedestrian trust and safety, for example, by adapting stopping profiles that will fit 'pedestrians' expectations in a specific situation. Additionally, in this study, we did not provide the pedestrian any information about the FAV reliability, nor did we investigate their behavior after getting familiar with the system. Future studies should investigate pedestrian behavior after they get familiar with the system. It will also be interesting to investigate the behavior in various levels of system reliability.

Pedestrians decided to cross (or not to cross) before the FAV completely stopped; therefore, they could not distinguish well between the two stopping profiles. Future studies should investigate which distance, speed, time to collision (TTC) are optimal for the 'pedestrians' Understanding, trust in FAV, and decision making.

This research used video-based scenarios, which simulate how pedestrians may behave in the FAV world in various crossing complexities and how the e-HMI influenced their decision. However, to extend these findings, it is necessary to conduct further studies in the real world and examine at first more complex crossing opportunities (e.g., more lanes, two-way streets, various car distances, speeds, etc.), as well as 'pedestrians' behavior with more than one FAVs simultaneity. Moreover, it is worth looking at adding sounds to the eHMI or an engine sound as those may help pedestrians estimate the FAV distance or increase awareness of its intentions.

Moreover, in future research, we should examine the effect of time pressure in real life. In addition, future studies should deeply explore how to increase Understanding of FAV's intentions in the two scenarios of close FAV and "Cross" suggestion and far FAV and "Don't cross" suggestion. It could be done by combining the FAV's behavior into various crossing complexities. In addition, more diverse, multicultural populations, such as the elderly and children, should be included in future research.

6. Acknowledgments

The first author, Michal Hochman, was funded by The Israeli Smart Transportation Research Center (ISTRC) via a doctoral student grant.

7. Reference

Ackermann, C., Beggiato, M., Schubert, S., & Krems, J. F. (2019). An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles? *Applied Ergonomics*, *75*(March 2018), 272–282. https://doi.org/10.1016/j.apergo.2018.11.002

Ackermans, S., Dey, D., Ruijten, P., Cuijpers, R. H., & Pfleging, B. (2020). *The Effects of Explicit Intention Communication , Conspicuous Sensors , and Pedestrian Attitude in Interactions with Automated Vehicles*. 1–14.

Clamann, M., Aubert, M., & Cummings, M. (2017a). *Evaluation of Vehicle-to-Pedestrian Communication Displays for Autonomous Vehicles*. *January*.

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

Deb, S., Carruth, D. W., & Strawderman, L. J. (2020). *A Survey Study to Explore Comprehension of Autonomous Vehicle' s Communication Features*. 1–12.

Deb, S., Strawderman, L. J., & Carruth, D. W. (2018). Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *59*, 135–149. https://doi.org/10.1016/j.trf.2018.08.016

Hochman, M., Parmet, Y., Oron-gilad, T., & Collet, C. (2020). Pedestrians' Understanding of a Fully Autonomous Vehicle' s Intent to Stop : A Learning Effect Over Time. *Frontiers in Psychology*, *11*(December), 1–11. https://doi.org/10.3389/fpsyg.2020.585280

Kaleefathullah, A. A., Merat, N., Lee, Y. M., Eisma, Y. B., Madigan, R., Garcia, J., & Winter, J. de. (2020). External Human–Machine Interfaces Can Be Misleading: An Examination of Trust Development and Misuse in a CAVE-Based Pedestrian Simulation Environment. *Human Factors*, *August*. https://doi.org/10.1177/0018720820970751

Liu, H., Hirayama, T., Morales, L. Y., & Murase, H. (2020). *What Is the Gaze Behavior of Pedestrians in Interactions with an Automated Vehicle When They Do Not Understand Its Intentions?* *May*. http://arxiv.org/abs/2001.01340

Lv, W., Lee, Y. M., Uzondu, C., & Madigan, R. (2021). *A distributed simulation study to investigate pedestrians' head- turning behaviour when crossing in response to automated and human-driven vehicles displaying different braking patterns*. *October*. https://doi.org/10.31234/osf.io/qt6ma

Mahadevan, K., Somanath, S., & Sharlin, E. (2018). Enabling Pedestrian Communication with Autonomous Vehicles. *CHI 2018 Workshop Interacting with Autonomous Vehicles: Learning from Other Domains*.

Ratchford, M., & Barnhart, M. (2012). Development and validation of the technology adoption propensity (TAP) index. *Journal of Business Research*, *65*, 1209–1215.

Robinette, P., Wagner, A., & Howard, A. (2013). Building and maintaining trust between humans and guidance robots in an emergency. *AAAI Spring Symposium, Stanford University*, 78–83. https://www.aaai.org/ocs/index.php/SSS/SSS13/paper/viewFile/5755/6007

Tabone, W., de Winter, J., Ackermann, C., Bärgman, J., Baumann, M., Deb, S., Emmenegger, C., Habibovic, A., Hagenzieker, M., Hancock, P. A., Happee, R., Krems, J., Lee, J. D., Martens, M., Merat, N., Norman, D., Sheridan, T. B., & Stanton, N. A. (2021). Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. *Transportation Research Interdisciplinary Perspectives*, *9*(December 2020), 100293. https://doi.org/10.1016/j.trip.2020.100293

Winter, J. De, & Shyrokau, B. (2019). *Conceptual Testing of Visual HMIs for Merging of Trucks Conceptual Testing of Visual HMIs for Merging of Trucks*. *April*.

Zuckerman, M. (2007). The sensation seeking scale V (SSS-V): Still reliable and valid. *Personality and Individual Differences*, *43*(5), 1303–1305. https://doi.org/10.1016/j.paid.2007.03.021