Implicit Interaction with an Autonomous Personal Mobility Vehicle: Relations of Pedestrians’ Gaze Behavior with Situation Awareness and Perceived Risks

Hailong Liu\textsuperscript{a,b}, Takatsugu Hirayama\textsuperscript{b,c}, Luis Yoichi Morales Saiki\textsuperscript{d}, and Hiroshi Murase\textsuperscript{b}

\textsuperscript{a}Graduate School of Science and Technology, Nara Institute of Science and Technology, Nara, Japan; \textsuperscript{b}Graduate School of Informatics, Nagoya University, Aichi, Japan; \textsuperscript{c}Faculty of Environmental Science, University of Human Environments, Aichi, Japan; \textsuperscript{d}Institute of Innovation for Future Society, Nagoya University, Aichi, Japan

\textbf{ABSTRACT}
Interactions between pedestrians and autonomous personal mobility vehicle (APMV) will increase with the popularity of autonomous driving systems. However, when the APMVs are applied in a mixed traffic environment after manual driving PMV (MPMV) have been popular, pedestrians may feel unsafe in the interactions when they are uncertain about the driving intention of the APMV. This study seeks to find a surrogate measure for pedestrians’ understanding of driving intention and perceived safety during the interaction with an APMV. We conducted an experiment to measure the gaze duration and subjective evaluations of the participants when they interacted with a PMV in manual and autonomous driving modes. Pedestrians fixed their gaze at the APMV longer when they did not accurately understand the driving intention than when they understood it. Furthermore, the pedestrians perceived danger when they did not clearly understand the driving intention of the APMV. Besides, these factors were different when pedestrians interact with an MPMV and an APMV.

1. Introduction

The functions and capabilities of autonomous driving systems (ADS) are continuously improving with increasing emphasis on safety. ADS is not only used in autonomous cars but also in personal mobility vehicles (PMV) (Andersen et al., 2016; Morales et al., 2018; Watanabe et al., 2015). The significant difference between a PMV and a car is its slow speed, and the PMV can be used in a traffic scenario that is expected to contain a mixture of it and pedestrians. As the ADS will be applied to PMVs, the autonomous PMVs (APMVs) may often interact with pedestrians in shared spaces, on sidewalks, and indoors, such as malls, airports and stations (Ali et al., 2019; Kobayashi et al., 2013; Morales et al., 2018). Thus, the number of interactions between APMVs and pedestrians will increase with the popularity of APMVs.

However, the APMVs may face the same dilemma as autonomous cars when the APMVs are applied in a mixed traffic environment after MPMVs have been popular. For example, one aspect of the resistance to popularization is that people do not trust autonomous cars (Pettigrew et al., 2019) when they do not understand the driving intention in interactions, particularly with unmanned car. We consider that this issue will also occur in the popularization of APMVs, or even more seriously because the APMVs have more opportunities to interact with pedestrians than the autonomous cars, and the interaction distance is closer. A typical example of such interaction is depicted in Figure 1. Pedestrians may feel worried when they want to pass in front of an APMV because it is difficult to determine whether the APMV has detected them and could keep them safe.

To solve this issue, if pedestrians confuse the driving intention of APMVs, then providing information to pedestrians regarding the driving intention is helpful to improve the pedestrians’ understanding and their perceived safety in the interactions (Habibovic et al., 2018; Rasouli & Tsotsos, 2020). Therefore, this paper focuses on the following research question:

\textbf{RQ:} How does the APMV evaluate whether pedestrians understand the driving intention and whether they feel safe in the interaction?

To answer this research question, the main aim of this paper is to find a surrogate measure to evaluate the uncertainty of pedestrians’ cognition and subjective feeling when they try to understand the driving intention of the APMV and when they feel safe in the interaction. We used a top-down approach to design an experiment on pedestrian–vehicle interaction based on a conceptual model explaining dynamic process of pedestrian’s decision-making. The experiment result shows the difference in pedestrians’ gaze durations and the difference in their subjective evaluations.
when they interact with the MPMV and the APMV; Besides, it verifies that the gaze duration of the pedestrians had significantly negative correlations with their understanding of driving intention and their perceived safety when they interact with the APMV.

2. Related works

2.1. Evaluation method of pedestrian–vehicle interactions

Many subjective evaluation methods have been used for pedestrian–vehicle interactions (de Clercq et al., 2019; Faas & Baumann, 2020; Y. M. Lee et al., 2019; Löcken et al., 2019; Watanabe et al., 2015). To establish communication between an APMV and a pedestrian, Watanabe et al. (2015) mounted an on-board projector to show the navigational intention of a robotic wheelchair. Then they investigated pedestrians’ understanding of the driving intention and feeling of comfort in the interaction by a questionnaire. Faas and Baumann (2020) used questionnaire and structured interview to evaluate the pedestrians’ understanding when an autonomous car used various types of external human-machine interface (eHMI) to show the yielding intention. de Clercq et al. (2019) asked participants to continuously evaluate their feeling of safety by pressing a button during the AV–pedestrian interaction. Similarly, Y. M. Lee et al. (2019) and Löcken et al. (2019) analyzed the pedestrians’ subjective evaluations by using a questionnaire when they interacted with an autonomous car in a virtual reality space.

Meanwhile, objective variables have been used to analyze the interactions between pedestrians and vehicles in a small number of studies (Dey et al., 2019; Fuest et al., 2018, 2020). Dey et al. (2019) conducted an eye-tracking study with 26 participants in a road-crossing situation. They found that as a manual driving vehicle was approaching, the gaze of the pedestrians gradually moved to the windshield at the driver’s position. In other words, as the vehicle approaches and the pedestrians cannot determine whether the vehicle will stop, they look at the driver to acquire the driver’s intention. However, for level 3–5 autonomous cars (SAE Technical Standards Board, 2016) as same as the APMV, pedestrians cannot understand the driving intention of the driver or the passenger because the autonomous cars usually do not require the operations of a driver. Therefore, the pedestrians can only recognize the driving intention of the APMVs from the implicit driving signals, such as speed, acceleration, and driving direction. Fuest et al. (2018) and Fuest et al. (2020) focused on the time for pedestrians to recognize an autonomous car’s intention in an interaction. To observe the time, participants were asked to press a button when they thought they recognized the driving intention. They compared the recognition time for the driving intention in different right-of-way situations, and it showed the differences in subjective evaluation for the understanding in various interaction situations. However, the above studies did not investigate the relation between the recognition time and the subjective understanding of the autonomous car’s driving intention.

2.2. Evaluation method of understanding

To answer the research question, surrogate measures of understanding must be discussed. Our initial idea referenced studies in cognitive psychology, education and human-computer interactions. A typical example is from Just and Carpenter (1976) who reported that a longer gaze duration indicated difficulty in extracting information. Similarly, Poole and Ball (2006) and Djamashi (2014) also focused on the gaze duration in the human–computer interactions because they considered that a longer gaze duration could be presented higher cognitive effort. Liversedge et al. (1998) reported that the gaze duration could be used to gain an understanding of the difficulty of text reading. In other words, if a text is more difficult to understand, then the reader will spend more time reading it. Similarly, Rayner et al. (2006) demonstrated that the processing time of understanding and the number of fixations increased when the difficulty of the text being reading increased. Besides, Martínez-Gómez and Aizawa (2014) reported that a combination of eye–movement features and text characteristics is useful to predict reader’s language skill and level of understanding. Furthermore, to predict the reader’s understanding of the text, Sanches et al. (2017) and Sanches et al. (2018) proved that the subjective understanding of the reader could be predicted accurately by using various features of the eye–gaze behavior, such as gaze duration.

For pedestrian–APMV interaction, we consider that pedestrians may continue to observe the APMV when they did not understand the intention of APMV because they may extract useful information difficulty from the APMV via implicit communication. Based on the above studies, the gaze duration of pedestrians could be considered as a surrogate measure of the pedestrians’ subjective understanding of APMV’s driving intention. pedestrians may continue to observe the APMV when they did not understand the intention of APMV because they extract useful information difficulty from the APMV via implicit communication. Thus,
our paper analyzes the relation between gaze duration and subjective feelings of pedestrians.

### 3. Aims and hypotheses

The main aim of this paper is to find a surrogate measure to evaluate the uncertainty of pedestrians’ cognition and subjective feeling when they try to understand the driving intention of the APMV and when they feel safe in the implicit interaction. Based on the related works, we found some clues about the relation between gaze duration and subjective feelings from related works. This relation should be further explained in the dynamic process of a pedestrian’s decision-making.

Thus, we proposed a conceptual model in our previous study (Liu et al., 2020), which is a dynamic model of the pedestrian’s decision-making process in an interaction with a vehicle. It is shown in Figure 2. This conceptual model includes three parts: situation awareness, hazard perception, and decision-making based on risk homeostasis.

This study focuses on situation awareness, which includes three steps: perception, comprehension, and projection (Endsley, 1995). Firstly, situation awareness relies on the perception of things in the surrounding environment, for example, APMV’s relative position, relative distance, and relative speed. Secondly, comprehension is taken as the understanding of the current state of the APMV, such as the driving intention of the APMV. Thirdly, the pedestrians will predict the driving behavior and moving trajectory of the APMV based on the result of comprehension. Based on the described process, we suppose that if the pedestrians do not understand the APMV’s driving intention clearly in the comprehension step, then they will continue to gaze at the vehicle until they believe they have correctly understood the driving intention of it.

After establishing situation awareness, the pedestrians realize hazards, such as anomaly detection, by comparing the predicted driving behavior of the APMV with their experience. Then, the subjective risk is evaluated by using the perceived hazards. Subsequently, the subjective risk could be seen as the degree of the pedestrian’s feeling of being threatened, that is, the perceived safety.

The pedestrians decide their behavior by comparing their subjective risk with their acceptable risk level according to the risk homeostasis theory (Wilde, 1982). If the perceived risk is lower than the acceptable risk level, the pedestrians will decide to engage in riskier behavior. In the opposite case, the pedestrians will decide to change their behavior and become more careful. It is also called risk compensation behavior. Note that, we consider that the gaze behavior can also be adjusted by the risk they already felt. For example, the high risk that pedestrians currently feel from the APMV may cause them to continue to observe the APMV at the next point in time to ensure their own safety.

Based on the proposed conceptual model, we formulated the following hypotheses:

- **H1**: If the gaze duration of a pedestrian on the APMV increases, then the pedestrian may not clearly understand the driving intention of the APMV;
- **H2**: If the pedestrian does not clearly understand the driving intention of the APMV, then the pedestrian may perceive danger.
- **H3**: If the pedestrian perceive danger, then their gaze duration on the APMV may increase.

In summary, we consider that the gaze duration could represent the pedestrians’ understanding of the driving intention of the APMV and their perceived safety in the interaction.

### 4. Method

We conducted a Wizard of Oz (WoZ) experiment: an experiment, in which the participants are led to believe they are interacting with an autonomous system, but the system is partially operated by an experimenter. Throughout our WoZ experiment, an APMV automatically drove on the pre-designed routes and the experimenter could secretly control the APMV to stop according to the actual situation during the participant-APMV interactions.

MPMVs are becoming popular as in recent years there has been an increase in their availability and usage (Shin et al., 2018; Zagorskas & Burinskienė, 2019). It is expected that these PMVs will gradually have partial to fully automated functionalities, i.e., APMVs. The experiments of this paper emulate the order in which MPMV and APMV are expected to spread out in the market. Therefore, we set up an experimental background imitating the popularization order of the MPMV and the APMV, that was the APMVs were just applied in the shared space after MPMVs have been popular. Thus, we asked participants to interact with an MPMV accumulating enough experience before interacting with the APMV.
According to the proposed model (Figure 2), we assumed that the gaze durations could be seen as a surrogate measure to evaluate the uncertainty of pedestrians’ cognition and subjective feeling. Thus, the goal of this experiment was to measure the gaze durations of the participants when they interacted with an MPMV and an APMV. Meanwhile, the participants’ subjective evaluations of the interactions were recorded using questionnaires.

This research complied with the American Psychological Association Code of Ethics and was approved by the Research Ethics Committee at the Institutes of Innovation for Future Society, Nagoya University (No. 2021-11). Informed consent was obtained from each participant.

### 4.1. Experimental condition

#### 4.1.1. Movement routes

The movement routes for pedestrian–PMV interactions are shown in Figure 3. Three driving routes were prepared in advance, which are assumed to have different levels of interaction complexity. We assume that the subjective risk perceived by participants from PMVs (i.e., MPMV and APMV) is higher in scenarios with higher interaction complexity. The participants were asked to walk at their normal pace from the start point to the goal point. Although the walking route of the participants was not restricted, the intended walking route could be assumed as the blue dot line in Figure 3.

Route 1 was considered to have a high interaction complexity. The interaction when the PMVs took Route 1 simulates a scene of crossing paths in an intended area in Figure 3 shown in yellow. Thus, the participants need to judge the driving direction and whether the PMVs will yield to them or not by continuous observation.

Route 2 is considered to have a medium interaction complexity. The intended interaction when the PMVs took Route 2 simulates a scene, in which the pedestrians and the PMVs approach and pass by each other on a straight road in the green area of Figure 3. In this scene, pedestrians only need to pay attention to the driving direction of the PMVs and avoid it.

Route 3 is considered to have a low interaction complexity. The interaction when the PMVs took Route 3 simulates a scene of a long-distance interaction between the pedestrians and the PMVs. The driving intention of the PMVs on Route 3 is clear because PMVs will not physically interact with the participants on this route, and the PMVs are no threat to the pedestrians. Note that although the PMVs do not approach the pedestrians on Route 3, the pedestrians would still need to observe, understand, and recognize the driving intention of the PMVs to determine whether their paths would cross.

In addition, we needed to confirm that factors unrelated to the interaction, such as participant’s curiosity, did not affect the gaze durations. In this case, we assumed that if the pedestrians’ gaze durations on the MPMV and the APMV were not significantly different when they interacted on Route 3, then the curiosity of pedestrians might have no obvious effect on the gaze durations.

To simulate the uncertainty of the scene where pedestrians interact with PMVs in real life, the sequence of driving routes including 21 trials was randomly selected before the experiment (1 → 1 → 2 → 1 → 3 → 2 → 1 → 3 → 2 → 1 → 1 → 3 → 1 → 2 → 1 → 1 → 1 → 2 → 1). It includes 11 trials of Route 1, 7 trials of Route 2, and 3 trials of Route 3. Besides, we take into account that none of the participants have experience of interacting with MPMV and APMV. In addition, we believe that participants’ situation model for MPMV and APMV needs to be improved in their interaction with PMVs in different situations (routes), which is a cumulative process. Combined with previous explanation of interaction complexity, all participants interact with the MPMV and the APMV by this sequence in order to reduce the effects of the route sequence.

#### 4.1.2. Vehicle

A PMV, WHILL Model CR, shown in the left part of Figure 4, was used as the experimental vehicle. It can be driven manually and automatically.

For manual driving, an operator rode on the MPMV and used the available joystick to manipulate it. The operator did not actively convey information about driving intention through action or language to the participant. Meanwhile, there were no turn lights or brake lights on the vehicle to represent the status of the vehicle to the participant.

For autonomous driving, the APMV was equipped with a multilayered lidar (Velodyne VLP-16) and wheel encoders. The lidar was utilized for self-localization on a previously built environmental map. The APMV had an automatic brake function that was applied when there was an obstacle within 0.5 meters directly in front of it. This function was used to ensure the safety of the participants in the experiment. The APMV could automatically drive following the pre-designed routes, but it did not have the functions to recognize the pedestrians and suitably interact with them, for example, it could not automatically yield the right of way.

![Figure 3. Movement routes.](image)
The experimenter used a wireless remote controller as shown in the top left part of Figure 4, to secretly control the APMV and stop it according to the actual situation during participant-APMV interactions. The maximum speed of the PMV was limited to 1 [m/s] for automated and manual driving. During the experiment, no devices were showing information about the driving status and intention of the APMV to the participant.

4.1.3. Participants

We recruited 24 experiment participants (17 males and 7 females) within the age range of 21–30 (mean: 24.4, standard deviation: 2.58) as pedestrians. They had different educational backgrounds because they came from various disciplines at university and various departments of companies. None of them had any prior experience of driving or interacting with the APMVs and autonomous cars.

To measure the gaze behavior of the participants during walking, they were asked to use a wearable eye tracker Tobii Pro Glasses 2 throughout the experiment. The field of view and gaze point from the participant’s perspective is presented in Figure 5. The red circle represents the participant’s gaze area and the red line represents the movement of the gaze point.

4.2. Experimental procedure

First, in order to reduce the participants’ gaze behavior caused by curiosity about the unique appearance of the PMV, we showed the participants the appearance of the PMV, including sensors, the body of the electric wheelchair, and the on-board computer.

Then, we explained to the participants the following information before the experiment:

1. Please walk at your normal pace from the start point to the goal point.
2. An experimental PMV will interact with you during your walk.
3. The PMV has three driving routes, and it will randomly select a route before each trial.
4. The PMV will be used for manual and fully autonomous driving.
5. An operator will ride on the PMV when driving manually, and no one will ride on the PMV when it is driving autonomously.
6. The maximum speed of the PMV is limited to 1 [m/s] (i.e., 3.6 [km/h]) for both manual and autonomous driving.
7. During autonomous driving, the APMV’s sensors can recognize the surrounding objects and you. As a result, the APMV will automatically determine its driving behavior when interacting with you. (False information)
8. There is no perfect system in the world, so this experiment still involves some risks.
9. If you think the behavior of the PMV threatens you, please stay away from it.
10. When the interaction is over, the experimenter will use a wireless remote controller to manually control the APMV, so it returns to its start position.

Next, the participants were requested to interact with the MPMV for 21 trials. Then, they were requested to interact with the APMV for 21 trials. This experimental sequence was designed to take into account the popularity order of MPMVs and APMVs because we considered that MPMVs are widely used earlier than APMVs from a realistic point of view. Therefore, we would like to allow participants to
experience the interaction with the APMV after fully experiencing the interaction with the MPMV.

It should be noted that the participants were not restricted to walking behaviors, in order to more realistically imitate life scenarios. Participants were only instructed to walk safely from the start point to the goal point. During this time, the participants were free to move around the experimental scene. For example, they can freely decide to walk or stay while interacting with PMV, and they can even choose to go back and detour.

After the completion of each trial of interaction, the participants were required to complete a questionnaire about their subjective evaluations. There were two evaluations on the questionnaire that were answered according to five-grade scales in Japanese. To evaluate the participant’s understanding or confusion related to the driving intention of the PMV during the interaction, the first evaluation question was the following with five-grade scale for answering:

Q1: How much do you think that you understood the intention of the PVM when you interacted with it?

U1: Completely did not understand;
U2: Did not understand much;
U3: Neutral;
U4: Mostly understood;
U5: Fully understood.

The second evaluation was used to evaluate the participant’s perceived safety during the interaction according to the following question with five-grade scale:

Q2: How much do you think that you perceived safety or danger during the interaction?

S1: Very dangerous;
S2: Slightly dangerous;
S3: Neutral;
S4: Slightly safe;
S5: Very safe.

Finally, an unstructured interview was conducted with each participant for 30–40 minutes after the interactions. The participants were mainly asked for their feelings and impressions of the interaction experiment.

In total, each participant spent about two hours participating in this experiment including introduction, contract, setting up equipment, experiment, interview, and a short-break after interacting with the MPMV and before interacting with the APMV.

4.3. Data preprocessing

In the experiment, the observed gaze data of participants #11 and #22 were excluded because their long eyelashes affected the collected gaze data which had a lot of defects. Besides, the gaze data from multiple trials were excluded due to the data saving problems and the gaze calibration problems. In total, gaze durations and subjective evaluations of 22 participants were measured in 447 trials with the MPMV (Route 1: 225 trials, Route 2: 154 trials, Route 3: 68 trials) and 455 trials with the APMV (Route 1: 235 trials, Route 2: 154 trials, Route 3: 66 trials).

Tobii Pro Glasses 2 measured the foreground video and sequence of gaze points of each participant. The method to calculate the gaze duration of each trial is shown in Figure 6. The size of the measured foreground video was $1920 \times 1080$ pixels. In each frame, the gaze point of the participant was recorded as a two-dimensional coordinate value on the plane of the foreground image. In this experiment, the central visual field of participants was defined as a circle. The gaze point was the center of a circle with a diameter of 108 pixels. If any part of the PMV area overlapped with part of the circle, then it was determined that the participant was gazing at the PMV. Under the above-mentioned conditions, the total time of the gaze at the PMV was calculated as the gaze duration in each trial.

5. Result

All trials where PMV drove on the Routes 1 and 2 had close interaction with participants except for the case where PMV drove on Route 3. When the PMV drove on Route 1, which simulated interaction at an intersection, the participants determined their walking behavior to the PMV depending on
In this subsection, we analyze whether pedestrians’ gaze durations on the MPMV and the APMV reported in the following subsections.

### 5.1. Gaze duration on the MPMV and the APMV

In this subsection, we analyze whether pedestrians’ gaze durations on MPMV and APMV driving on different routes are different when they interact with the APMV after having been familiar with interacting with the MPMV. Thus, their gaze durations on the MPMV and the APMV driving on the three routes were compared by using a non-parametric two-way repeated-measures ANOVA with post-hoc comparisons.

The gaze durations of the 22 participants are presented by box plots in Figure 7. The light gray and dark gray boxes show the gaze durations on the MPMV and the APMV, respectively. These median values, interquartile ranges (IQR) and the results of Shapiro–Wilk test for normality are shown in Table 1. The Shapiro–Wilk test rejected that those gaze durations were sampled from the Gaussian distribution. Besides, the result of homogeneity of variance test via Levene test (Levene, 1960) rejected that those gaze durations had the same variance ($W = 13.63, p < 0.001$).

Based on the above test results, a non-parametric two-way repeated-measures analysis of variance (ANOVA) with post-hoc comparisons was used to compare the gaze durations of the participants when interacting with the MPMV and the APMV on different routes. Meanwhile, we also considered that the trials (the amount of data) on three routes were unbalanced. Therefore, a two-way aligned rank transformed ANOVA (two-way ART ANOVA) (Wobbrock et al., 2011) based on a linear mixed-effects model was conducted on the influence of two effects, i.e., driving modes and routes, on the gaze durations.

The result in Table 2 showed that all effects were statistically significant at the 0.05 significance level. The main effect for routes yielded an $F$ ratio of $F = 141.56, p < 0.0001$. Besides, the main effect for driving modes also yielded an $F$ ratio of $F = 2.9, p < 0.0001$. Moreover, the interaction effect was significant, $F = 9.43, p = 0.0001$. The post-hoc ART comparisons (Elkin et al., 2021) with Bonferroni adjustment method were conducted, the results of which are shown in Table 3. This result would provide information about gaze durations within each effect were significantly different ($p < 0.05$) except for (Route 1, MPMV) – (Route 2, APMV) ($p = 1.0$); (Route 2, MPMV) – (Route 3, APMV) ($p = 0.2585$); (Route 2, MPMV) – Route 3, MPMV ($p = 1.0$); and (Route 3, APMV) – (Route 3, MPMV) ($p = 0.1413$).

### 5.2. Subjective evaluations for interacting with the MPMV and the APMV

This section shows the results of pedestrians’ subjective evaluations, i.e., understanding of driving intention and the perceived safety, for the MPMV and the APMV driving on different routes. Because of that their subjective evaluations were reported by themselves through the five-grade scale, the ART ANOVA based on a linear mixed-effects model was conducted for analyzing the influence of driving modes and routes on those discrete subjective evaluation data.

#### 5.2.1. Understanding of driving intention

The subjective evaluation result for the understanding of driving intention is shown in Figure 8. The median scores and IQRs of participant evaluation scores for the understanding of driving intention were shown in Table 4.

Table 5 shows the influence of two effects of driving modes and routes on the evaluation scores for the understanding of driving intention by using a two-way ART ANOVA based on a linear mixed-effects model. All effects were statistically significant at the 0.05 significance level. In specific, the main effect for routes yielded an $F$ ratio of $F = 128.84, p < 0.0001$. Besides, the main effect for driving modes also yielded an $F$ ratio of $F = 70.69, p < 0.0001$. Moreover, the interaction effect was significant, $F = 11.03, p < 0.0001$.

The results of a post-hoc ART comparisons with Bonferroni adjustment method are shown in Table 6. These results would present that the evaluation scores for the understanding of driving intention within each effect were

---

**Figure 7.** Gaze duration on the MPMV and the APMV driving on the three routes.

---

**Table 1.** Data distribution and Shapiro–Wilk test for gaze durations.

| Data distribution | Shapiro–Wilk test for normality |
|-------------------|-------------------------------|
|                   | Median | IQR  | $w$ Value | $p$ Value |
| Route 1 MPMV      | 2.86   | 2.21 | 0.90      | <0.0001   |
| Route 1 APMV      | 3.94   | 3.25 | 0.96      | <0.0001   |
| Route 2 MPMV      | 1.38   | 1.47 | 0.95      | <0.0001   |
| Route 2 APMV      | 3.08   | 2.27 | 0.94      | <0.0001   |
| Route 3 MPMV      | 1.19   | 1.76 | 0.88      | <0.0001   |
| Route 3 APMV      | 1.78   | 2.59 | 0.93      | <0.0001   |

**Table 2.** Analysis of variance of aligned rank transformed gaze duration data by linear mixed-effects model.

| Effects            | df res | $F$ value | $Pr(>F)$ |
|--------------------|--------|-----------|-----------|
| Routes             | 2      | 875.06    | <0.0001   |
| Driving modes      | 1      | 875.02    | <0.0001   |
| Routes, Driving modes | 2 | 875.03    | 0.0001    |

Type III Wald $F$ tests with Kenward–Roger df.
Degrees-of-freedom method: Kenward–Roger; p value adjustment: Bonferroni method.

Figure 8. Accumulative bar diagram showing the ratios of evaluation scores for the understanding of driving intention of the MPMV and the APMV. (The ratios are rounded to the zero decimal place.)

Table 3. Post-hoc comparisons of aligned rank transformed gaze duration data based on the linear mixed-effects model.

| Contrast                        | Estimate | SE    | df  | t Ratio | p Value |
|---------------------------------|----------|-------|-----|---------|---------|
| (Route 1, APMV) – (Route 1, MPMV) | 110.69   | 15.58 | 875.05 | 7.62 | <0.0001 |
| (Route 1, APMV) – (Route 2, APMV) | 92.14    | 17.31 | 875.01 | 5.32 | <0.0001 |
| (Route 1, APMV) – (Route 2, MPMV) | 321.57   | 17.31 | 875.01 | 18.58 | <0.0001 |
| (Route 1, APMV) – (Route 3, APMV) | 262.96   | 23.26 | 875.01 | 11.31 | <0.0001 |
| (Route 1, APMV) – (Route 3, MPMV) | 338.06   | 23.01 | 875.09 | 14.69 | <0.0001 |
| (Route 1, MPMV) – (Route 2, APMV) | –26.55  | 17.47 | 875.06 | –1.52 | 1.0000 |
| (Route 1, MPMV) – (Route 2, MPMV) | 202.88   | 17.47 | 875.06 | 11.61 | <0.0001 |
| (Route 1, MPMV) – (Route 3, APMV) | 144.27   | 23.38 | 875.03 | 6.17  | <0.0001 |
| (Route 1, MPMV) – (Route 3, MPMV) | 219.37   | 23.14 | 875.14 | 9.48  | <0.0001 |
| (Route 2, APMV) – (Route 2, MPMV) | 229.43   | 19.03 | 875.00 | 12.06 | <0.0001 |
| (Route 2, APMV) – (Route 3, APMV) | 170.82   | 24.56 | 875.00 | 6.95  | <0.0001 |
| (Route 2, APMV) – (Route 3, MPMV) | 245.92   | 24.32 | 875.05 | 10.11 | <0.0001 |
| (Route 2, MPMV) – (Route 3, APMV) | –58.61   | 24.56 | 875.00 | –2.39 | 0.2585 |
| (Route 2, MPMV) – (Route 3, MPMV) | 16.49    | 24.32 | 875.05 | 0.68  | 1.0000 |
| (Route 3, APMV) – (Route 3, MPMV) | 75.10    | 28.86 | 875.03 | 2.60  | 0.1413 |

Table 4. Data distribution of the subjective evaluation result for the understanding of driving intention.

| Route | Median | IQP |
|-------|--------|-----|
| Route 1 MPMV | U₄     | 1   |
|Route 1 APMV | U₄     | 2   |
|Route 2 MPMV | U₄     | 1   |
|Route 2 APMV | U₄     | 0   |
|Route 3 MPMV | U₄     | 0   |
|Route 3 APMV | U₄     | 0   |

Table 5. Analysis of variance of aligned rank transformed evaluation scores for the understanding of driving intention by linear mixed-effects model.

| Effects                  | F     | df | df res  | Pr(>F) |
|--------------------------|-------|----|---------|--------|
| Routes                   | 128.84| 2  | 875.21  | <0.0001|
| Driving modes            | 70.69 | 1  | 875.07  | <0.0001|
| Routes: Driving modes    | 11.03 | 2  | 875.12  | <0.0001|

Type III Wald F tests with Kenward–Roger df.

significantly different (p<0.05) except for (Route 1, MPMV) – (Route 2, APMV) (p=0.0917) and (Route 3, APMV) – (Route 3, MPMV) (p=1.0).

5.2.2. Perceived safety

Figure 9 shows the subjective evaluation result for the perceived safety. The median scores and IQRs of participant evaluation scores for interaction with the MPMV and the APMV driving on the three routes were shown in Table 7. A two-way ART ANOVA based on a linear mixed-effects model was used to show the influence of two effects of driving modes and routes on the evaluation scores for the perceived safety. The result in Table 8 shows that all effects were statistically significant at the 0.05 significance level. In specific, the main effect for routes yielded an F ratio of F=117.03, p<0.0001. Besides, the main effect for driving modes also yielded an F ratio of F=34.78, p<0.0001. Moreover, the interaction effect was significant, F=6.73, p=0.0013.

Then, as shown in Table 9, a post-hoc ART comparisons with Bonferroni adjustment method was used to present that the significantly different level of the evaluation scores for the perceived safety within each effect was p<0.05, except for (Route 1, APMV) – (Route 2, APMV) (p=1.0); (Route 1, MPMV) – (Route 2, APMV) (p=0.0668); (Route 1, MPMV) – (Route 2, MPMV) (p=0.1939); and (Route 3, APMV) – (Route 3, MPMV) (p=1.0).

5.3. Relation between the gaze duration and the understanding of driving intention

According to the proposed conceptual model (Figure 2), if pedestrians do not understand the driving intention of the APMV in an interaction, they will continue to observe the APMV behavior. To verify Hypothesis H₄, a correlation analysis was used to show the relation between the gaze duration and the understanding of the driving intention in this section.
Table 6. Post-hoc comparisons of aligned rank transformed evaluation scores for the understanding of driving intention based on the linear mixed-effects model.

| Contrast                     | Estimate | SE  | df  | t Ratio | p Value |
|------------------------------|----------|-----|-----|---------|---------|
| (Route 1, APMV) – (Route 1, MPMV) | −131.70  | 18.15 | 875.16 | −7.26  | <0.0001 |
| (Route 1, APMV) – (Route 2, APMV) | −75.76  | 20.17 | 875.05 | −3.76  | 0.0028  |
| (Route 1, APMV) – (Route 2, MPMV) | −224.89 | 20.17 | 875.05 | −11.15 | <0.0001 |
| (Route 1, APMV) – (Route 3, APMV) | −327.32 | 27.10 | 875.03 | −12.08 | <0.0001 |
| (Route 1, APMV) – (Route 3, MPMV) | −357.51 | 26.81 | 875.30 | −13.33 | <0.0001 |
| (Route 1, MPMV) – (Route 2, APMV) | 55.95   | 20.36 | 875.20 | 2.75   | 0.0917  |
| (Route 1, MPMV) – (Route 2, MPMV) | −93.19  | 20.36 | 875.20 | −4.58  | 0.0001  |
| (Route 1, MPMV) – (Route 3, APMV) | −195.61 | 27.24 | 875.11 | −7.18  | <0.0001 |
| (Route 1, MPMV) – (Route 3, MPMV) | −225.81 | 26.96 | 875.48 | −8.38  | <0.0001 |
| (Route 2, APMV) – (Route 3, APMV) | −251.56 | 28.62 | 875.00 | −8.79  | <0.0001 |
| (Route 2, APMV) – (Route 3, MPMV) | −281.75 | 28.34 | 875.17 | −9.94  | <0.0001 |
| (Route 2, MPMV) – (Route 3, APMV) | −102.42 | 28.62 | 875.00 | −3.58  | 0.0055  |
| (Route 2, MPMV) – (Route 3, MPMV) | −132.61 | 28.34 | 875.17 | −4.68  | <0.0001 |
| (Route 3, APMV) – (Route 3, MPMV) | −30.19  | 33.62 | 875.12 | −0.90  | 1.0000  |

Degrees-of-freedom method: Kenward–Roger; p value adjustment: Bonferroni method.

Figure 9. Accumulative bar diagram showing the ratios of evaluation scores for the perceived safety interacting with the MPMV and the APMV. (The ratios are rounded to the zero decimal place.)

Table 7. Data distribution of the subjective evaluation result for the perceived safety.

| Route   | Median | IQP |
|---------|--------|-----|
| Route 1 | MPMV   | $S_4$ | 2   |
|         | APMV   | $S_4$ | 2   |
| Route 2 | MPMV   | $S_4$ | 1   |
|         | APMV   | $S_5$ | 0.75|
| Route 3 | MPMV   | $S_5$ | 0.92|
|         | APMV   | $S_5$ | 0   |

Table 8. Analysis of variance of aligned rank transformed evaluation scores for the perceived safety by linear mixed-effects model.

| Effects | $F$    | df  | df res | Pr(>|F|) |
|---------|--------|-----|--------|---------|
| Routes  | 117.03 | 2   | 875.09 | <.0001  |
| Driving modes | 34.78 | 1   | 875.03 | <.0001  |
| Routes: Driving modes | 6.73 | 2   | 875.04 | 0.0013  |

Type III Wald $F$ tests with Kenward–Roger df.

The gaze durations on the MPMV and the APMV were respectively divided into five groups according to the five scores of subjective evaluations. The results are shown in Figures 10 and 11. Note that the horizontal axes represent the five-grade scales of evaluation and the vertical axes represent the gaze duration. The mean values are shown by white triangles in each of the figures. A linear regression was used to visualize the correlation between the two variables. The solid black line represents the result of the linear regression. The gray area around the solid black line represents the 95% confidence interval for the regression. The larger the gray area, the less accurate the linear regression is and vice versa.

Figure 10 illustrates the relation between the participants’ subjective evaluation for their understanding of driving intention and their gaze duration in their interactions with the MPMV. The result of Spearman’s correlation analysis showed that there was a significant weak negative correlation between the two variables ($r_s = −0.22, p = 1.73 \times 10^{-6}$).

For the interaction with the MPMV, significance test were used to analyze the differences among gaze durations which corresponds with five subjective evaluation scores of the understanding of driving intention. Figure 10 shows that Shapiro–Wilk test did not reject that the gaze duration corresponding with all scores were sampled from the Gaussian distribution except $U_4$ ($W = 0.95, p < 0.001$) and $U_5$ ($W = 0.92, p < 0.001$).

Moreover, Levene test did not reject that the distributions of their five groups of gaze durations had the same variance ($W = 0.21, p = 0.934$). Based on the result of Shapiro–Wilk test, Kruskal–Wallis H test (Kruskal & Wallis, 1952), which is a non-parametric one-way ANOVA, was used to verify the difference in gaze duration corresponding to each scale. The result of Kruskal–Wallis H test rejected that the gaze durations on MPMV corresponding with five evaluation
scores were sampled from a same population \((H = 25.45, p < 0.001)\), i.e., there were significant differences in gaze duration among the scales when the participants interacted with the MPMV. After that, the multiple comparison test via Dwass-Steel-Critchlow-Fligner (DSCF) test (Douglas & Michael, 1991) was used to verify the significant difference in gaze duration between each pair of scales. The results of the DSCF test showed that there were significant differences in gaze duration on the MPMV between the following pairs: \(U_2 - U_5\) \((p < 0.05)\) and \(U_4 - U_5\) \((p < 0.01)\).

For the case of the AP MV, Figure 11 shows the relation between the participants’ subjective evaluation for their understanding of driving intention and their gaze duration in their interactions with the APMV. The Spearman’s correlation coefficient between the two variables was \(r_s = -0.36\) with \(p = 3.66 \times 10^{-15}\).

Besides, Figure 11 shows that Shapiro–Wilk test did not reject that the gaze duration on AP MV corresponding with all scores were sampled from the Gaussian distribution except \(U_7\) \((W = 0.94, p < 0.01)\) and \(U_6\) \((W = 0.88, p < 0.001)\). Besides, Levene test did not reject that the distributions of their five groups of gaze durations had the same variance \((W = 1.32, p = 0.260)\). Then, a non-parametric one-way ANOVA Kruskal–Wallis H test rejected that the gaze durations on AP MV corresponding with five evaluation scores were sampled from the same population \((H = 64.71, p < 0.001)\). Moreover, the results of the DSCF test showed that there were significant differences in gaze duration on the AP MV between the following pairs: \(U_2 - U_4\) \((p < 0.01)\), \(U_1 - U_5\) \((p < 0.001)\), \(U_2 - U_5\) \((p < 0.001)\), \(U_3 - U_5\) \((p < 0.001)\) and \(U_4 - U_5\) \((p < 0.001)\).
5.4. Relation between the understanding of driving intention and the perceived safety

According to the proposed conceptual model (Figure 2), the perceived risk of pedestrians will increase, that is, the perceived safety will decrease, when they do not understand the driving intention of the PMV. To verify Hypothesis $H_2$, we investigated the relation between the two subjective evaluations, the understanding of driving intention and the perceived safety.

Considering that these two subjective evaluations are discrete data, conditional probability of each scale for the perceived safety $P(S_i|U_j)$ was calculated. Here, $S_i \in \{S_1, S_2, S_3, S_4, S_5\}$ and $U_j \in \{U_1, U_2, U_3, U_4, U_5\}$. Specifically,

$$P(S_i|U_j) = \frac{P(S_i \land U_j)}{P(U_j)}, \quad (1)$$

$$P(U_j) = \frac{\text{count}(U_j)}{N}, \quad (2)$$

where $N$ is the total number of trials, the count $(\cdot)$ is a function for counting the selected times. If $P(S_i|U_j)$ is high, then it could be explained by $U_j$ having a high probability leading to $S_i$.

For the subjective evaluation result of the MPMV, the conditional probabilities $P(S_i|U_j)$ are shown in Figure 12. $P(S_1|U_1)$ and $P(S_2|U_1)$ were equal and the highest for the choice of $U_1$. Besides, $P(S_4|U_3)$ was the highest but it was also very similar to the $P(S_3|U_3)$. Meanwhile, $P(S_2|U_2)$, $P(S_4|U_4)$ and $P(S_5|U_5)$ were the highest conditional probabilities when the participants chose $U_2$, $U_4$, and $U_5$, respectively. Moreover, Spearman’s correlation analysis showed that there was a significantly strong correlation.
between those two subjective evaluation results for the MPMV ($r_s = 0.61, p = 3.69 \times 10^{-16}$).

Figure 13 shows the conditional probabilities $P (S_i | U_j)$ when the participants interacted with the APMV. $P (S_1 | U_1)$ and $P (S_4 | U_1)$ had a high probability when $U_1$ was chosen. For $U_2$, the highest conditional probability was $P (S_1 | U_2)$, but $P (S_1 | U_2)$ was the second highest. $P (S_1 | U_1)$ and $P (S_4 | U_2)$ were the first and second highest when the $U_3$ was chosen. $P (S_2 | U_3)$, and $P (S_3 | U_3)$ were higher than the other conditional probabilities. Besides, a significantly strong correlation between those two subjective evaluation results for the APMV ($r_s = 0.70, p = 2.50 \times 10^{-68}$) was analyzed by Spearman’s correlation.

### 5.5. Relation between the gaze duration and the perceived safety

In the above, we clarified the relation between the gaze duration and the understanding of driving intention, and the relation between the understanding of driving intention and the perceived safety. Refer to the proposed conceptual model (Figure 2) and Hypothesis $H_3$, the high risk that pedestrians currently feel from APMV may cause them to continue to observe the APMV at the next point in time to ensure their own safety. On the other hand, the low perceived risk may cause pedestrians take a risky gaze behavior, i.e., do not continue to observe APMV. To test this hypothesis, the relation between them was also investigated by the correlation analysis and ANOVA with post-hoc comparisons.

The gaze durations on the MPMV and the APMV were respectively divided into five groups according to the five evaluation scores of the perceived safety, i.e., $S_1$ to $S_5$. For the interaction with the MPMV, the relation between the participants’ perceived safety and their gaze duration on the MPMV is shown in Figure 14. There is a significant weak negative correlation ($r_s = -0.21, p = 5.67 \times 10^{-6}$) between them via Spearman’s correlation analysis.

An ANOVA with post-hoc comparisons was used to analyze the differences among gaze durations which correspond with five subjective evaluation scores of the perceived safety. Before the ANOVA, the distribution of the gaze durations should be tested first. Figure 14 shows that Shapiro–Wilk test did not reject that the gaze durations corresponding with all scores were sampled from a same population ($W = 0.96, p < 0.001$) and $S_4$ ($W = 0.93, p < 0.001$). Besides, Levene test rejected that the distributions of their five groups of gaze durations had the same variance ($W = 3.46, p < 0.01$). Therefore, Kruskal–Wallis H test, which is a non-parametric one-way ANOVA, was used to verify the difference in gaze durations corresponding to each scale of the perceived safety. The result of Kruskal–Wallis H test rejected that the gaze durations on MPMV corresponding with five evaluation scores were sampled from a same population ($H = 33.58, p < 0.001$). After that, the multiple comparison test via the DSCF test was used to verify the significant difference in gaze duration between each pair of scales. The results of the DSCF test showed that there were significant differences in gaze duration on the MPMV between the following pairs: $S_2 - S_3$ ($p < 0.05$) and $S_4 - S_5$ ($p < 0.001$).

Besides, Figure 15 shows the relation between the participants’ perceived safety and their gaze duration on the APMV. Spearman’s correlation coefficient between these two variables was $r_s = -0.40$ with $p = 3.59 \times 10^{-19}$, i.e., a significant negative correlation.
Moreover, Figure 15 shows that Shapiro–Wilk test did not reject that gaze duration on APMV corresponding with all scores were sampled from the Gaussian distribution except $S_3 (W = 0.97, p < 0.05)$, $S_4 (W = 0.97, p < 0.01)$ and $S_5 (W = 0.91, p < 0.001)$. Besides, Levene test did not reject that the distributions of their five groups of gaze durations had the same variance ($W = 3.01, p < 0.05$). Then, a non-parametric one-way ANOVA Kruskal–Wallis H test rejected that the gaze durations on APMV corresponding with five evaluation scores of the perceived safety were sampled from a same population ($H = 77.46, p < 0.001$). Moreover, the results of the DSCF test showed that there were significant differences in gaze duration on the APMV between the following pairs: $S_2 - S_3 (p < 0.001)$, $S_3 - S_5 (p < 0.001)$ and $S_4 - S_5 (p < 0.001)$.

6. Discussion

6.1. Gaze duration on the MPMV and the APMV

Referring to Table 2, for the gaze durations, the main effects for routes and driving modes as well as their interaction effect were significant. Meanwhile, Figure 7 shows interaction effect that the gaze durations on APVM were longer than that on MPMV with the increase of interaction complexity. This result illustrates that the gaze durations of participants on PMVs were not only effected by driving modes, but also effected by interaction complexities (routes, see Section 4.1.1.).

Figure 7 and Table 3 show that there was not a significant difference between the gaze durations on the MPMV and the APMV driving on the Route 3 ((Route 3, APMV) – (Route 3, MPMV); $p = 0.1413$). We would like to note that Route 3 was designed as a route which can make pedestrians clearly and quickly understand the driving intention of the PMV without feeling in danger (see Section 4.1.1.). The result of Route 3 precisely indicates that the participants did not significantly increase their gaze durations on the APMV due to their curiosity about the appearance of the APMV. From another perspective, when the participants interacted with PVMs on Route 1 and 2, the effect of their curiosity on the subjective evaluations for the MPMV and the APMV could be excluded.

Regarding the reasons for the difference in the subjective evaluations of participants between the MPMV and the APMV driven on routes 1 and 2, we considered that there were two aspects. The first aspect could be considered that it is difficult for the participants to obtain information through implicit communication, quickly establishing a correct and trust-able mental model (Endsley, 2000) to support their situation model for interacting with the APMV, e.g., did not understand APMV’s behavioral logic and function limitations. This might make participants suspicious and under-confident of their predicted driving intentions of APMV lacking a driver. Moreover, this incomprehension of APMV’s driving intentions had led to an increase in the perceived risk of the participants. The second aspect could be considered as the trust in the human operator on the MPMV, i.e., in the interview, the participants reported that they were not sure whether the APMV would guarantee their safety as same as the human operator of the MPMV.

6.2. Subjective evaluations for interacting with the MPMV and the APMV

From Figures 8 and 9 as well as Tables 6 and 9, these results mentioned above showed that the 22 participants thought that the driving intentions of the APMV were more significantly difficult to understand than the driving intention of the MPMV on routes 1 and 2. In addition, their perceived safety was also significantly lower when they interacted with the APMV than when they interacted with the MPMV on routes 1 and 2, which matches the result in Merat et al. (2018). However, when the MPMV and the APMV drove on Route 3, the participants did not differ significantly in their subjective evaluations.

Note that, Routes 1 and 2 were designed as close interactions which might make the participants feel more danger or more confused about the PMV’s driving intention than on Route 3, i.e., Route 3 was designed as a long-distance interaction with low risk. As same as the result of participants’ gaze durations shown in Figure 7, the result of Route 3 precisely indicates again that the participants did not significantly increase their subjective evaluations on the APMV due to their curiosity about the appearance of the APMV. From another perspective, when the participants interacted with PVMs on Route 1 and 2, the effect of their curiosity on the subjective evaluations for the MPMV and the APMV could be excluded.

6.3. Relation between the gaze duration and the understanding of driving intention

Figure 10 shows that the participants’ gaze durations on the MPMV had significant differences between $U_4 - U_5$ and between $U_4 - U_5$ but there were no significant differences among $U_1$ to $U_4$. Meanwhile, Figure 11 shows that the gaze durations on the APMV for $U_5$ was significantly lower than gaze durations for the other scores. The above shows that the participants’ gaze durations on the MPMV reduced significantly only when they completely understood its driving intention.

Moreover, Figure 10 shows that Spearman’s correlation coefficient of the $-0.22$ between the participants’ understanding of the MPMV’s driving intention and their gaze durations. Compared with this result, there was a stronger negative correlation for interacting with the APMV, which was $-0.36$ as shown in Figure 11. Although the Spearman’s correlations for the MPMV and the APMV were not very strong, the above results suggested...
that the participants’ gaze durations on the PMV were related to their understanding of the driving intention, especially for the APMV. Therefore, Hypothesis $H_1$ could be considered correct, especially when pedestrians interact with APMV.

This finding is in line with previous studies (Just & Carpenter, 1976; Liversedge et al., 1998; Rayner et al., 2006; Sanches et al., 2017, 2018), i.e., gaze duration could be used to indicate the difficulty in extracting information, understanding and cognitive workload. Moreover, some studies about human-computer interactions also reported that the longer gaze duration present higher cognitive effort (Djamasbi, 2014; Poole & Ball, 2006). We considered that when the participants did not understand the intentions of PMVs, they would try to make a good effort of cognition for understanding it.

6.4. Relation between the understanding of driving intention and the perceived safety

The above results indicate that the participants’ different understanding of driving intention would produce correspondingly different perceptions of safety, that is, the more the participants understood the driving intention of the MPMV and the APMV, the higher their perceived safety of the driving behavior. This conclusion is in line with Cao et al. (2021) who reported that perceived safety was correlated with the understanding of driving intentions. Similarly, previous studies (de Clercq et al., 2019; Habibovic et al., 2018) also commented that making pedestrians understand an autonomous car’s intentions could help them improve their perceived safety.

This result also verified that pedestrians’ have a high probability of effect on their perceived subjective risk in our proposed conceptual model (Figure 2). This further illustrates the importance of APMV and MPMV making pedestrians to understand their driving intention in an interaction. Therefore, Hypothesis $H_2$ is not only correct for pedestrians interacting with the APMV but also correct for them interacting with the MPMV.

6.5. Relation between the gaze duration and the perceived safety

Figures 14 and 15 show that Spearman’s correlation coefficient between the participants’ gaze durations and their perceived safety when they interacted with the MPMV and the APMV were $-0.21$ and $-0.40$. That is, the gaze duration of participants on the APMV gradually increased when the perceived safety they felt from the APMV gradually reduced. Besides, the post-hoc comparison results in Figures 14 and 15 showed that if the participants felt very safe from the MPMV or APMV, their gaze durations on it reduced significantly. In other words, the participants were not likely to continue to gaze at the PMVs in a situation deemed safe by them. Moreover, this result could also show that the gaze duration of the participants on the PMVs could reflect their perceived safety, i.e., Hypothesis $H_3$ is correct.

6.6. Trust factors in the interaction

In the interviews after the experiment, a portion of the participants reported that they did not pay much attention to the MPMV because the operator would avoid them even if they did not understand the driving intention. Therefore, the interaction may be affected by the participants’ trust in the operator riding on the MPMV. We found that the difference in participants’ trust in the MPMV and the APMV may lead to a variety of differences in their gaze behaviors (gaze durations) and in their subjective evaluations for the MPMV and the APMV. For example, we found a trend through Figures 10 and 11 that the gaze duration on the APMV was gradually greater than that on the MPMV as the understanding of driving intention reduced, except when $U_1$ was chosen, i.e., they did not completely understand the driving intention of the MPMV. That is to say, even when the participants did not understand the driving intention ($U_1$ and $U_2$) of the MPMV, their gaze duration did not increase much. Referring to the results of the perceived safety in Figures 14 and 15, the same trend can be noticed. The reason for this trend is speculated that the participants might have trusted the operator who manually drove the MPMV even when they did not understand the driving intentions ($U_1$ and $U_2$) and did not feel safe ($S_1$ and $S_2$). For the APMV, we speculated that the participants’ trust in the APMV might be lower than their trust in the MPMV. Thus, as their understanding of driving intention became ambiguous and their perceived safety decreased, their trust in the APMV might have declined, while they might have increased their observation of the APMV to gain more information to protect themselves from danger.

Based on the above discussions, an important factor for pedestrian–PMV interactions could be considered the pedestrians’ potential trust. This may be the key factor for the APMVs to achieve popularity in society.

6.7. Effect on the eHMI

In the experiment, we found that it was difficult for participants to understand APMV’s driving intentions through implicit communication. For this issue, eHMI could be seen as an effective communication method for conveying driving intention of an autonomous car to pedestrians (Burns et al., 2019; Dey et al., 2021; Keferböck & Rien, 2015). We consider that eHMI is also important in the interaction between pedestrians and APMVs in order to help them negotiate the right-of-way because the APMVs have more opportunities to interact with pedestrians than the autonomous cars, and the interaction distance is closer. However, Li et al. (2018) and J. Lee et al. (2021) reported that eHMI might cause pedestrians to be distracted and careless without comprehensive understanding of traffic situations.

In our view, APMVs as autonomous cars need to communicate with pedestrians at an appropriate timing, e.g., a timing when pedestrians want to obtain information about driving intention from it. Therefore, based on this study, APMVs could predict the pedestrian’s understanding of its
driving behavior in real time by recognizing the gaze durations on it, thereby improving the perceived safety when pedestrians interact with APMVs in the future.

6.8. Effect on the autonomous driving systems

In the field of autonomous driving systems, a lot of studies attempted to predict the intentions of pedestrians to help autonomous vehicles generate their driving behaviors such as (Abughalieh & Alawneh, 2020; Fang et al., 2017; Quintero et al., 2017). These studies are all to ensure the physical safety (no collision) of autonomous cars interacting with pedestrians. Few autonomous driving systems generated the driving behaviors taking into account the subjective feelings of pedestrians. For example, Anthony et al. (2020) proposed an on-board system that could recognize the gaze behavior of pedestrians to determine whether pedestrians were aware of the autonomous car when they interacted with each other.

Based on our study, perhaps in the future autonomous cars as same as APMVs can take into account the subjective feelings of pedestrians to plan and generate driving behavior. Then, the public acceptance of autonomous cars and APMVs will increase.

6.9. Limitations of the experimental results

A limitation of this study is that the experimental results can be applied to PMVs interacting with pedestrians at low speed, but this result may be difficult to explain an interaction with high-speed PMVs such as Segway. The reason is that the gaze duration of the pedestrians may depend on the speed of PMVs.

Besides, the random selection resulted in unequal selection times for each route in order to simulate the interactions in real-world scenarios. This unequal selection will be addressed in our future studies while there was no clear date to suggest that this had an effect on the findings in this paper.

In addition, the behavior of PMVs was safe for the participants, as recommended by the research ethics committee. This led to the overwhelming majority of participants’ evaluations of PVMs clustered in positive evaluations, such as $U_4$, $U_5$, $S_4$, and $S_5$. Therefore, the correlation coefficient between gaze duration and the two subjective ratings were not very high, although they were significant.

Moreover, the participants in this experiment were young people because it is difficult for us to invite elderly people to participate in our experiment under the influence of COVID-19.

Furthermore, to verify our hypothesis in a controlled environment, we conducted the experiment in a simple environment without other pedestrians and vehicles and without environmental noise. In the actual environment, other objects may attract the attention of pedestrians which may reduce the gaze duration on the APMV.

7. Conclusions

7.1. Summary

In order to increase the popularity of APMVs within society, providing pedestrians with timely motion cues for understanding the driving intention and improving their perceived safety are considered effective. Thus, the main purpose of this paper was to find a surrogate measure to represent the understanding of driving intention and the perceived safety by pedestrians in their interactions with APMVs. For this purpose, we accorded to a conceptual model proposed in our previous study which shows a dynamic process of a pedestrian’s decision-making interacting with a PMV. Based on this conceptual model, the pedestrian–PMV interaction experiment was implemented to verify that the gaze duration of pedestrians could be used to represent their understanding of driving intention and their perceived safety in their interactions with the APMV.

We found the following from experimental results:

1. Participants’ gaze durations on the APMV were significantly longer than their gaze durations on the MPMV in a close interaction.
2. Participants evaluated that the driving intention of the APMV was more significantly difficult to understand than that of the MPMV and the perceived safety for the APMV was also significantly lower than that for the MPMV.
3. Participants’ perceived safety relied on their understanding of the driving intention of the MPMV and the APMV.
4. Participants’ gaze durations had a correlation with those two subjective evaluations. Especially when the participants interacted with the APMV, their gaze durations reduced significantly when they felt very safe and completely understand the driving intention of it.

Recalling the research question “RQ: How does the APMV evaluate whether pedestrians understand the driving intention and whether they feel safe in the interaction?”, our answer is that the gaze duration of the pedestrians could be seen as a surrogate measure for evaluating their understanding of the driving intention of the APMV and their perceived safety in a close interaction.

7.2. Future works

We will use the same experimental method to study the interaction between pedestrians of a wider age and a high-speed PMV or an autonomous car in a more complex environment. Besides, we will attempt to build a system for APMVs, which will recognize pedestrians’ gaze behavior to predict their understanding of the driving intentions. Based on the design concept proposed by Li et al. (2021) and the driving behavior visualization method proposed by Liu et al. (2017), an eHMI on APMV that can quickly, clearly, and timely convey driving intentions to pedestrians will also be developed, thus improving the acceptability of APMVs.
addition, the results of this study indicate that pedestrians’ trust level in APMVs may influence their gaze duration. Some related works have also pointed out that trust is an important factor for human-vehicle interactions (Holländer et al., 2019; Liu & Hiraoka, 2019; Liu et al., 2021; McAllister et al., 2017). In future studies, it will be necessary to discuss the trust calibration for pedestrians interacting with APMVs.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Funding
This work was supported by JSPS KAKENHI Grant [Numbers 20K19846, 19K12080], Japan and JST COI Grant [Number JPMJCE1317], Japan.

ORCID
Hailong Liu http://orcid.org/0000-0003-2195-3380

References
Abughali, K. M., & Alawneh, S. G. (2020). Predicting pedestrian intention to cross the road. IEEE Access, 8, 72558–72569. https://doi.org/10.1109/ACCESS.2020.2987777
Ali, S., Lam, A., Fukuda, H., Kobayashi, Y., Kuno, Y. (2019). Smart wheelchair maneuvering among people. In D. S. Huang, Z. K. Huang, & A. Hussain (Eds.), Intelligent computing methodologies. ICIC 2019. Lecture notes in computer science (pp. 32–42). Springer.
Andersen, H., Eng, Y. H., Leong, W. K., Zhang, C., Kong, H. X., Pendleton, S., Ang, M. H., & Rus, D. (2016). Autonomous personal mobility scooter for multi-class mobility-on-demand service [Paper presentation]. IEEE International Conference on Intelligent Transportation Systems, Rio de Janeiro, Brazil (pp. 1753–1760).
Anthony, S. E., Misra, K., & Faller, A. W. (2020). System and method of predicting human interaction with vehicles. https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2019112912
Burns, C. G., Oliveira, L., Thomas, P., Iyer, S., Birrell, S. (2019). Pedestrian decision-making responses to external human-machine interface designs for autonomous vehicles [Paper presentation]. IEEE Intelligent Vehicles Symposium (pp. 70–75).
Cao, J., Lin, L., Zhang, J., Zhang, L., Wang, Y., & Wang, J. (2021). The development and validation of the perceived safety of intelligent connected vehicles scale. Accident Analysis and Prevention, 154, 106912. https://doi.org/10.1016/j.aap.2021.106912
de Clercq, K., Dietrich, A., Velasco, J. P. N., de Winter, J., & Happee, R. (2019). External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. Human Factors, 61(8), 1353–1370. https://doi.org/10.1177/0018720819836343
Dey, D., Matviienko, A., Berger, M., Pfleleging, B., Martens, M., & Terken, J. (2021). Communicating the intention of an automated vehicle to pedestrians: The contributions of eHMI and vehicle behavior. it - Information Technology, 63(2), 123–141. https://doi.org/10.1515/iti-2020-0025.
Dey, D., Walker, F., Martens, M., & Terken, J. (2019). Gaze patterns in pedestrian interaction with vehicles: towards effective design of external human-machine interfaces for automated vehicles [Paper presentation]. Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 369–378).
Djamal, S. (2014). Eye tracking and web experience. AIS Transactions on Human-Computer Interaction, 6(2), 37–54. https://doi.org/10.17705/1hchi.00060
Douglas, C. E., & Michael, F. A. (1991). On distribution-free multiple comparisons in the one-way analysis of variance. Communications in Statistics - Theory and Methods, 20(1), 127–139. https://doi.org/10.1080/03610929108830487
Elkin, L. A., Kay, M., Higgins, J. J., & Wobbrock, J. O. (2021). An aligned rank transform procedure for multifactor contrast tests [Paper presentation]. arXiv Preprint, arXiv:2102.11824. https://doi.org/10.1145/3472474.3474784
Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. Human Factors, 37(1), 32–64. https://doi.org/10.1518/00187209579049543
Endsley, M. R. (2000). Situation models: An avenue to the modeling of mental models [Paper presentation]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 44, pp. 61–64). https://doi.org/10.1177/154193120004400117
Faas, S. M., & Baumann, M. (2020). Yielding light signal evaluation for self-driving vehicle and pedestrian interaction. In T. Ahram, W. Karwowski, S. Pickl, & R. Tair (Eds.), Human systems engineering and design II (pp. 189–194). Springer International Publishing.
Fang, Z., Vázquez, D., & López, A. M. (2017). On-board detection of pedestrian intentions. Sensors, 17(10), 2193. https://doi.org/10.3390/s17102193
Fuest, T., Maier, A. S., Bellem, H., & Bengler, K. (2020). How should an automated vehicle communicate its intention to a pedestrian? – A virtual reality study. In T. Ahram, W. Karwowski, S. Pickl, & R. Tair (Eds.), Human systems engineering and design II (pp. 195–201). Springer International Publishing.
Fuest, T., Michalowski, L., Träarris, L., Bellem, H., & Bengler, K. (2018). Using the driving behavior of an automated vehicle to communicate intentions – A wizard of oz study [Paper presentation]. IEEE International Conference on Intelligent Transportation Systems (pp. 3596–3601).
Habibovic, A., Lundgren, V. M., Andersson, J., Klingegård, M., Lagström, T., Sirkka, A., Fagerlönë, J., Edgren, C., Fredriksson, R., Krupenia, S., Saluå, D., & Larsson, P. (2018). Communicating intent of automated vehicles to pedestrians. Frontiers in Psychology, 9, 1336. https://doi.org/10.3389/fpsyg.2018.01336
Holländer, K., Wintersberger, P., & Butz, A. (2019). Overtrust in external cues of automated vehicles: An experimental investigation [Paper presentation]. Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 211–221).
Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. Cognitive Psychology, 8(4), 441–480. https://doi.org/10.1016/0010-0285(76)90015-3
Kefereb, F., & Rieners, A. (2015). Strategies for negotiation between autonomous vehicles and pedestrians. In A. Weisbecker, M. Burmester, & A. Schmidt (Eds.), Mensch und computer 2015 – workshopband (pp. 525–532). De Gruyter.
Kobayashi, Y., Suzuki, R., Sato, Y., Ara, M., Kuno, Y., Yamazaki, A., & Yamazaki, K. (2013). Robotic wheelchair easy to move and communicate with companions [Paper presentation]. Chi ’13 Extended Abstracts on Human Factors in Computing Systems (pp. 3079–3082). Association for Computing Machinery. https://doi.org/10.1145/2468356.2479615
Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion analysis of variance. Journal of the American Statistical Association, 47(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441
Lee, J., Daimon, T., & Kitazaki, S. (2021). Negative effect of external human-machine interfaces in automated vehicles on pedestrian crossing behaviour: A virtual reality experiment. In N. L. Black, W. P. Neumann, & I. Noy (Eds.), Human Factors in Computing Systems, vol. 221 (pp. 3079–3082). Association for Computing Machinery. https://doi.org/10.1145/3468356.3479615
Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), Contributions to probability and statistics (pp. 278–292). Stanford University Press.
Li, Y., Cheng, H., Zeng, Z., Liu, H., & Sester, M. (2021). Autonomous vehicles drive into shared spaces: ethno design concept focusing on vulnerable road users [Paper presentation]. 2021 IEEE International Intelligent Transportation Systems Conference (ITSC) (pp. 1729–1736). https://doi.org/10.1109/ITSC48978.2021.9564515

Li, Y., Dikmen, M., Hussein, T. G., Wang, Y., & Burns, C. (2018). To cross or not to cross: Urgency-based external warning displays on autonomous vehicles to improve pedestrian crossing safety [Paper presentation]. Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 188–197).

Liu, H., Hirayama, T., & Watanabe, M. (2021). Driving behavior model considering driver’s overtrust in driving automation system [Paper presentation]. Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings (pp. 115–119).

Liu, H., Hirayama, T., Morales, L. Y., & Murase, H. (2020). What timing for an automated vehicle to make pedestrians understand its driving intentions for improving their perception of safety? [Paper presentation]. 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC) (pp. 1–6).

Liu, H., Hirayama, T., & Watanabe, M. (2021). Importance of instruction for pedestrian-automated driving vehicle interaction with an external human machine interface: Effects on pedestrians’ situation awareness, trust, perceived risks and decision making [Paper presentation]. 2021 IEEE Intelligent Vehicles Symposium (IV) (pp. 748–754). https://doi.org/10.1109/IV48863.2021.9575246

Liu, H., Taniguchi, T., Tanaka, Y., Takenaka, K., & Bando, T. (2017). Visualization of driving behavior based on hidden feature extraction by using deep learning. IEEE Transactions on Intelligent Transportation Systems, 18(9), 2477–2489. https://doi.org/10.1109/TITS.2017.2649541

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Chapter 3 – Eye movements and measures of reading time. In G. Underwood (Ed.), Eye guidance in reading and scene perception (pp. 55–75). Elsevier Science Ltd.

Löcken, A., Golling, C., & Rienar, A. (2019). How should automated vehicles interact with pedestrians?: A comparative analysis of interaction concepts in virtual reality [Paper presentation]. Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 262–274). ACM. https://doi.org/10.1145/3342197.3344544

Martínez-Gómez, P., & Aizawa, A. (2014). Recognition of understanding level and language skill using measurements of reading behavior [Paper presentation]. Proceedings of the 19th International Conference on Intelligent User Interfaces (pp. 95–104). ACM. https://doi.org/10.1145/2557500.2557546

SAE Technical Standards Board. (2016). J3016: Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles (pp. 1–30). SAE International.

McAllister, R., Gal, Y., Kendall, A., van der Wilk, M., Shah, A., Cipolla, R., & Weller, A. (2017). Concrete problems for autonomous vehicle safety: Advantages of bayesian deep learning [Paper presentation]. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17 (pp. 4745–4753). https://doi.org/10.24963/ijcai.2017/661

Merat, N., Louw, T., Madigan, R., Wilbrink, M., & Schieben, A. (2018). What externally presented information do VRUs require when interacting with fully automated road transport systems in shared space? Accident: Analysis and Prevention, 118, 244–252. https://doi.org/10.1016/j.aap.2018.03.018

Morales, Y., Akai, N., & Murase, H. (2018). Personal mobility vehicle automated navigation through pedestrian flow: A data driven approach for parameter extraction [Paper presentation]. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 3438–3444).

Petigrew, S., Worrall, C., Talati, Z., Fritschi, L., & Norman, R. (2019). Dimensions of attitudes to autonomous vehicles. Urban, Planning and Transport Research, 7(1), 19–33. https://doi.org/10.1080/21650020.2019.1604155

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In C. Ghaoui (Ed.), ENCYCLOPEDIA of human computer interaction (pp. 211–219). IGI Global.

Quintero, R., Parra, I., Lorenzo, J., Fernández-Llorca, D., & Sotelo, M. (2017). Pedestrian intention recognition by means of a hidden Markov model and body language [Paper presentation]. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC) (pp. 1–7). https://doi.org/10.1109/ITSC.2017.8317766

Rasouli, A., & Tsotsos, J. K. (2020). Autonomous vehicles that interact with pedestrians: A survey of theory and practice. IEEE Transactions on Intelligent Transportation Systems, 21(3), 900–918. https://doi.org/10.1109/TITS.2019.2901817

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. Scientific Studies of Reading, 10(3), 241–255. https://doi.org/10.1207/s15327979ssr1003_3

Sanches, C. L., Augereau, O., & Kise, K. (2017). Using the eye gaze to predict document reading subjective understanding [Paper presentation]. 2017 14th Iapr International Conference on Document Analysis and Recognition (ICDAR), (Vol. 08, pp. 28–31). https://doi.org/10.1109/ICDAR.2017.377

Sanches, C. L., Augereau, O., & Kise, K. (2018). Estimation of reading subjective understanding based on eye gaze analysis. PLoS One, 13(10), e0206213. https://doi.org/10.1371/journal.pone.0206213

Shin, G. W., Lee, K.-J., Park, D., Lee, J. H., & Yun, M. H. (2018). Personal mobility device and user experience: A state-of-the-art literature review. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 62(1), 1336–1337. https://doi.org/10.1177/154193118621303

Watanabe, A., Ikeda, T., Morales, Y., Shinozawa, K., Miyashita, T., Hagita, N. (2015). Communicating robotic navigational intentions [Paper presentation]. IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 5763–5769).

Wilde, G. J. (1982). The theory of risk homeostasis: Implications for safety and health. Risk Analysis, 2(4), 209–225. https://doi.org/10.1111/j.1539-6924.1982.tb01384.x

Wobbrock, J. O., Findlater, L., Gergle, D., & Haggins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures [Paper presentation]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 143–146). https://doi.org/10.1145/1978942.1978963

Zagorskas, J., & Burinskiiené, M. (2019). Challenges caused by increased use of e-powered personal mobility vehicles in European cities. Sustainability, 12(1), 273. https://doi.org/10.3390/su12010273

About the Authors

Hailong Liu is an Assistant Professor at Graduate School of Science and Technology, NAIST, Japan. He received his Ph.D. degrees in Engineering from Ritsumeikan University, Japan in 2018. His research interests include machine learning and deep learning to analyze driving behavior and over-trust, eHMI, motion sickness in human-autonomous vehicle interactions.

Takatsugu Hirayama is a Professor at Faculty of Environmental Science, University of Human Environments and a Visiting Professor at Graduate School of Informatics, Nagoya University, Japan. He received his Ph.D. degrees in Engineering Science from Osaka University in 2005. His research interests include computer vision and human-computer interaction.

Luis Yoichi Morales Saiki is a Visiting Associate Professor at the Institutes of Innovation for Future Society, Nagoya University, Japan. He received his Ph.D. degrees in Robotics from University of Tsukuba in 2009. His research interests include autonomous navigation, spatial cognition, perception, and environment modeling.

Hirosi Murase is a Professor Emeritus and Designated Professor at Graduate School of Informatics, Nagoya University, Japan. He received his Ph.D. degrees in Electrical Engineering from Nagoya University in 1987. His research interests include computer vision and intelligent vehicle.