Causal Discovery from Psychological States to Walking Behaviors for Pedestrians Interacting an APMV Equipped with eHMIs

Hailong Liu[1,*], Yang Li[2], Toshihiro Hiraoka[3] and Takahiro Wada[1]

Abstract— **This study aims to investigate the causal relationships from pedestrians' psychological states to their walking behavior during interactions with an autonomous personal mobility vehicle (APMV) featuring automation capabilities ranging from SAE levels 3 to 5. A subjective experiment was conducted, where various external human-machine interfaces (eHMIs) were designed to induce participants to experience different levels of subjective feelings and generate corresponding walking behaviors. By employing a structural equation model named DirectLiNGAM to analyze the collected data for causal discovery, the results of causal discovery align with the hypothesized model of the pedestrian's cognition-decision-behavior process. Furthermore, the experimental results have enriched the detailed causal relationships within the hypothesized model, i. e., the outcomes of situation awareness lead to a sense of danger, trust in APMV and a sense of relief; the outcomes of situation awareness, the sense of danger and trust in APMV lead to hesitation in decision-making; and the outcomes of situation awareness, the sense of danger and hesitation lead to walking behaviors.**

I. INTRODUCTION

Autonomous personal mobility vehicles (APMVs) represent a category of small autonomous vehicles (AVs) featuring automation capabilities ranging from SAE levels 3 to 5 [1]. APMVs are specifically designed for short-distance mobility, and are suitable for anyone, without being restricted to use only by the elderly or individuals with disabilities [2]. Due to their slow speed and low acceleration, APMVs are suitable for use in shared spaces with other traffic participants, such as sidewalks [3], shopping centers, stations, and school campuses [4], [5]. Consequently, APMVs will inevitably engage with other traffic participants, including vulnerable road users like pedestrians. During the interaction between pedestrians and the APMV, Liu et al. reported that pedestrians were inclined to perceive danger due to confusion in understanding the driving intentions of the APMV [5], [6]. Addressing this challenge, equipping APMV with an external human-machine interface (eHMI) was reported as an effective strategy to effectively communicate the vehicles' driving intentions to pedestrians [2], [4].

Currently, most research on eHMI is focused on the interaction between humans and AVs [7]–[12]. The aforementioned studies draw conclusions regarding the impact on pedestrians' psychological states and behaviors by comparing various experimental conditions, such as different eHMIs, driving behavior conditions, and so forth. For example, clear information from the eHMI could help pedestrians understand the vehicle's intention [9], [12], enhanced sense of safety [8], [9], allows pedestrians to trust the AV [9], [12], reduce the decision-making time and hesitation [7], [9] and improve the crossing initial time (CIT) [10], [11].

There is a shortage of studies discussing the processes from pedestrians' psychological states to walking behaviors during interactions with AVs and using eHMIs. Specifically, there is a lack of research on this process during interactions between pedestrians and APMVs.

To address this research gap, the purpose of this study is to explore the causal relationships between psychological states and walking behaviors of pedestrians when interacting with APMVs equipped with eHMI. Additionally, this study aims to validate the previously proposed hypothesized model based on the causal exploration results.

II. CAUSAL DISCOVERY

A. Hypothesized Model of Pedestrian's Cognition-Decision-Behavior Process

To discover causal relations from psychological states to walking behaviors of pedestrians interacting with the APMV, this study is based on a hypothesized model of the pedestrian's cognition-decision-behavior process, as proposed by [5], [6], [9].

As presented in Fig. 1, the hypothesized model comprises four key components, i. e., situation awareness [13], risk homeostasis [14], decision-making and behavior generation. In the part of situation awareness,it describes a cognitive process of pedestrians which involves perception, comprehension, and projection, constituting their overall awareness of the surrounding environment. Thereafter in the risk homeostasis process, pedestrians perceive hazards based on prediction results and evaluate the subjective risk (e. g., , the sense of danger, the sense of relief), in the current situation by considering their own level of risk acceptance (i.e., target risk). Additionally, Liu et al. considered that pedestrians' trust in the APMV interacts with the target risk, influencing subjective risk evaluation [9]. Afterward, pedestrians decide on behaviors, such as walking, by comparing subjective risk with the acceptable risk level (i.e., target risk). Following this decision-making process, the body executes specific walking behaviors. As they engage in these walking behaviors, interactions with objects in the surrounding environment occur, perpetuating the aforementioned process in a continuous loop.

[1]Hailong Liu and Takahiro Wada are with the Graduate School of Science and Technology, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan.

[2]Yang Li is with Institute of Human and Industrial Engineering, Karlsruhe Institute of Technology, Engler-Bunte-Ring 4, Karlsruhe, 76133, Germany.

[2]Toshihiro Hiraoka is with Japan Automobile Research Institute, 1-1-30, Shibadaimon, Minato-ku, Tokyo 105-0012, Japan.

*CONTACT Hailong Liu. ✉ : liu.hailong@is.naist.jp

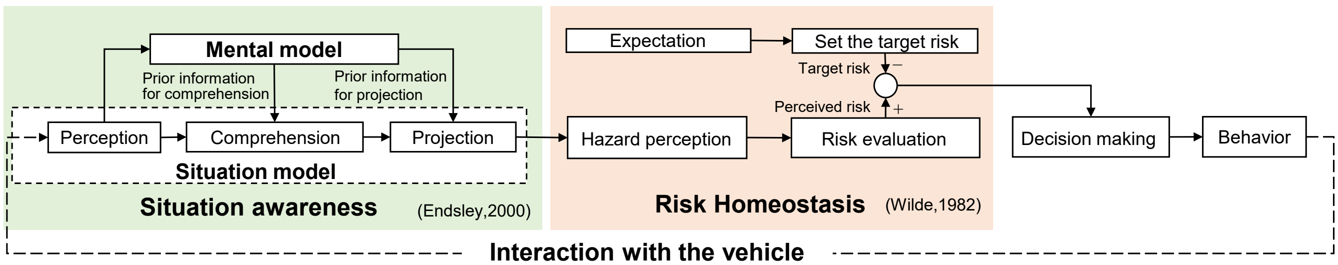

Fig. 1: A hypothesized model of pedestrian's cognition-decision-behavior process in human-vehicle interactions [5], [6], [9].

Importantly, it should be noted that pedestrians' behaviors will dynamically influence their next moment's situation awareness through interactions with the surrounding environment, as depicted by the dashed line in Fig. 1. This indirect effect, i.e., indirect causal relationship, is not within the scope of discussion in this study. Exactly for this reason, we consider that the direct causal process, i.e., generation process, from pedestrians' perception in situation awareness to behavior is unidirectional and non-cyclical. It can be adequately represented by a directed acyclic graph (DAG).

B. *Causal Discovery via Direct Linear Non–Gaussian Acyclic Model (DirectLiNGAM)*

The purpose of causal discovery is to estimate the data generation process of among observed variables. Specifically, it is to identify the causal relations among observed variables, determine which variables have direct or indirect causal influences on others within observed data. Structural equation model (SEM) has been extensively employed to examine causal relations among observed variables. In which, Shimizu et al. proposed a linear non–gaussian acyclic model (LiNGAM) for estimating DAG-based SEM by using the non-Gaussianity of disturbance variables [15]. The LiNGAM can be presented by

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e} \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ represents the observed variables, $\boldsymbol{e} \in \mathbb{R}^n$ represents the non-Gaussian independent disturbance variables, and $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ denotes a adjacency matrix. The n is the number of variables. Within matrix $\boldsymbol{A}$, $a_{i,j}$ represents the connection strength of the directed edge from x_i to x_j, illustrating the causal relation between them.

Generally, the objective of causal discovery by SEM is to find an optimal matrix $\boldsymbol{A}$ that describes the generative process, i. e., the causal relationships, among the observed data. Since LiNGAM assumed that the causal process can be represented by a DAG, it permutes the matrix $\boldsymbol{A}$ to a strictly lower triangular matrix by simultaneous equal row and column permutations. The lower triangular matrix $\boldsymbol{A}$ could be estimated by the independent component analysis (ICA) [15] but it may depend on the initially guessed state or the choice of parameters, making it challenging to guarantee that LiNGAM will converge to the correct solution in a finite number of steps [16]. To address this issue, Shimizu et al. proposed DirectLiNGAM which can directly extract causal structures for LiNGAM [16]. DirectLiNGAM infers exogenous variables by calculating the mutual information between variables and thier residuals. Once an exogenous variable is determined, it is excluded, and the inference process continues for the remaining variables to identify additional exogenous variables. Through terative repetition of these steps, the causal order among variables is estimated. After determining the causal order, DirectLiNGAM uses a least squares regression to estimate the adjacency matrix A with a strictly lower triangular structure based on the estimated causal order. For specific algorithms of DirectLiNGAM, please refer to [16].

In this paper, we assume that the generation process from psychological states to walking behaviors of pedestrians is unidirectional and non-cyclical, meaning it can be represented by a DAG. Therefore, DirectLiNGAM is used for this study.

III. EXPERIMENT

To discover the causal relationship between pedestrians' psychological states and behaviors during the interaction with an APMV, an interaction experiment was conducted. This experiment has been carried out with the approval of the Research Ethics Committee of Nara Institute of Science and Technology (No. 2022-I-55).

A. *Participants*

We invited 18 participants (Ages: 22 to 38 years, Avg.: 28 years, Std.: 5.5 years). All participants did not have previous experience of interacting with APMV. The total duration of the experiment was approximately one hour. Each participant received 1,000 Japanese Yen as a reward.

B. *Autonomous Personal Mobility Vehicle and eHMI*

A WHILL Model CR robotic wheelchair with an autonomous driving system was used as the APMV (see Fig. 2). This APMV is equipped with a Velodyne VLP-16 LiDAR and a PC controls for autonomous driving on a pre-designed route.

A screen with a speaker was mounted on top of the APMV, serving as an eHMI. The eHMI can exhibit driving intentions by displaying corresponding text on the screen and vocalizing them through its speaker.

C. *Experimental Site*

We set up a street-crossing scene, as shown in Fig. 3, in an indoor space of 10 m$\times$10 m. The pedestrian path is a total of 7.5 meters, including 4.5 meters from the starting point to the intersection, a street crossing width of 1.5 meters, and 1.5 meters from the street crossing to the exit. Additionally, APMV's path is a total of 8.25 meters, including 4.75 meters from the starting point to the intersection, a street crossing width of 2 meters, and 1.5 meters from the street crossing to the parking area. A baffle plate is set up next to the starting point of the participants. The baffle plate is designed to obstruct the participants' line of sight, ensuring that they can see the APMV from a consistent position only after they start walking.

D. Driving Behaviors of the APMV

The velocity profile of the APMV is depicted by the red line in Fig.3. The APMV initiates acceleration from a standstill at a distance of 4.75 meters from the intersection stop line. After covering a distance of 1.5 meters, it reaches a speed of 1.6 m/s, maintains this speed for 1.75 meters, and then initiates deceleration, ultimately coming to a stop after moving an additional 1.5 meters. In all trials, the APMV's velocity profile remains consistent, regardless of variations in eHMI conditions.

E. eHMI Conditions

To obtain a variety of psychological and behavioral data from pedestrians during the interaction with APMV, we implemented four distinct eHMI conditions. These conditions were designed to observe participants' diverse levels of comprehension and projection regarding the APMV's driving intentions and their impact on subjective feelings and walking behavior.

1) No eHMI: The eHMI neither presents information on the screen nor emits any sound, thereby providing no assistance to pedestrians in comprehending and anticipating the driving intentions of the APMV.

2) Early eHMI: The APMV, moving at a constant high speed, displays "I will stop" on the eHMI screen and provides a voice cue to say the message when it is 3.25 meters away from the stop line (see Fig. 3). We assume that participants may find it easier to predict the driving intentions of the APMV under this condition. Nevertheless, comprehension of the driving intentions at that moment might still pose a challenge due to the eHMI prompt occurring before the APMV decelerates.

3) Sync eHMI: At a distance of 1.5 meters from the stop line, the eHMI issues visual and voice cues, stating "I will stop", synchronously as the APMV initiates deceleration (see Fig. 3). The sync eHMI was assumed that it can enhance participants' understanding and prediction of the APMV's driving intentions.

4) Late eHMI: After the APMV comes to a stop before the intersection stop line, the eHMI will prompt "I stopped" using both visual and voice cues (see Fig. 3). The visual cue will continue until the APMV departs after the pedestrian completes crossing the road. As the eHMI offers no information before the APMV stops, we assume it aids pedestrians in understanding the current driving intentions but cannot assist in predicting the APMV's driving behaviors before it stops.

Fig. 2: The designed eHMI on the APMV can convey driving intentions to pedestrians through visual and voice cues.

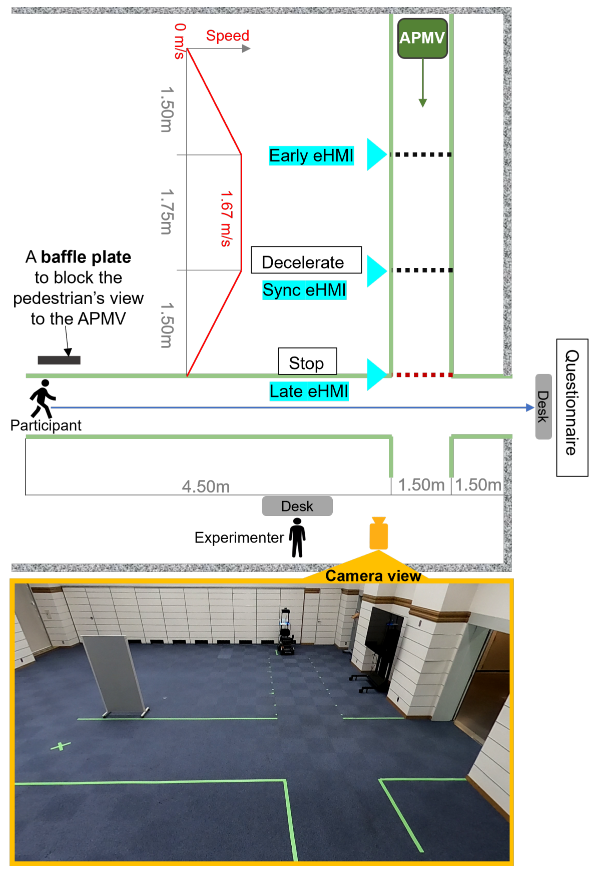

Fig. 3: Experimental settings

TABLE I. Experiential sequences of eHMI experience for the 18 participants.

Participants	Order of experience with eHMI conditions			
(N=18)	1st eHMI	2nd eHMI	3rd eHMI	4th eHMI
N=3	No eHMI →	Early eHMI →	Sync eHMI →	Late eHMI
N=3	No eHMI →	Early eHMI →	Late eHMI →	Sync eHMI
N=3	No eHMI →	Sync eHMI →	Early eHMI →	Late eHMI
N=3	No eHMI →	Sync eHMI →	Late eHMI →	Early eHMI
N=3	No eHMI →	Late eHMI →	Early eHMI →	Sync eHMI
N=3	No eHMI →	Late eHMI →	Sync eHMI →	Early eHMI

F. Experimental Procedure

First, participants received an introduction to the experiment, which included information on the hardware of the APMV and its autonomous driving system. Then, information about the eHMI, including displayed contents and voices, was introduced through examples. However, the triggering conditions and timing for activating the eHMI were not introduced. After the participants had no further concerns regarding the introduction of the experiment, they signed the informed consent and began the experiment.

In the experiment, each participant was pseudo-randomly assigned to one of six experiential sequences (see Table I) in order to minimize the potential impact of sequence effects on experimental results. In addition, considering that all participants lack prior interaction experience with the APMV, we positioned the *No eHMI* condition at the beginning of the sequence. The remaining three eHMI conditions provide six potential sequence combinations. As detailed in Table I, each of these six experience sequences will be undergone by three participants. Note that each eHMI condition will be carried out continuously for three interactive trials to obtain stable subjective evaluations and behavioral data.

In each interactive trial, participants were instructed to walk from the starting point (see the cross mark on Fig.3) to exit through a doorway, which included passing through an intersection. During this process, participants are instructed to walk naturally and decide whether the APMV should be yielded to—whether to stop and allow the APMV to cross or not, as would be done in real traffic. At the end of each trial, participants were required to complete a post-trial questionnaire at the desk located outside the door. Once the questionnaire was completed, participants were instructed to return to the starting point and repeat the rest of the trials.

G. Measurements

1) Post-trial Questioners: Referring to [9], following each interaction trial, pedestrians were requested to provide subjective feedback by answering six questions using a 5-point Likert scale, i. e., 1="strongly disagree", 2="disagree", 3="neutral", 4="agree", and 5="strongly agree", representing interval scales. These six questions are:

- Q1: It was easy to understand the driving intentions of the APMV.
- Q2: It was easy to predict the driving behaviors of the APMV.
- Q3: I felt it was dangerous to cross the road when I encountered the APMV.
- Q4: I trusted the APMV to interact with me safely when I crossed the road.
- Q5: I felt a sense of relief when I crossed the road.
- Q6: I felt hesitant to make the decision of crossing the road or not when I encountered the APMV.

Q1 and Q2 are used to assess the comprehension and projection steps within pedestrians' situation awareness. Simultaneously, Q3, Q4, and Q5 are applied to evaluate the risk homeostasis process. Furthermore, Q6 is employed to assess hesitation in decision-making, specifically gauging the difficulty of deciding to cross the road.

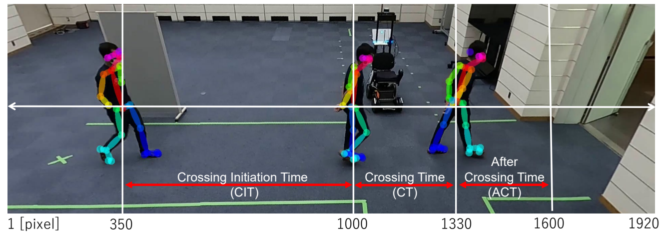

Fig. 4: The position of participants are estimated using OpenPose with BODY 25 joint set. The participants' MidHip keypoint is used to calculate the crossing initiation time (CIT), the crossing time (CT) and after crossing time (ACT).

2) Walking times: As shown in Fig. 3, the walking times of the participants in the three stages are calculated. **Crossing Initiation Time (CIT)**: the time it takes for participants to move from the location where APMV is visible to the beginning of entering the intersection is observed. **Crossing Time (CT)**: the time spent crossing the intersection. **After Crossing Time (ACT)**: the time spent from crossing the intersection to reaching the endpoint.

To accurately measure the above three walking times, as shown in Fig. 3, the entire process of participants crossing the road is captured by a camera (1920×1080 pixels with 30 FPS). OpenPose [17] based on the BODY 25 joint set is used to extract skeletal features of participants during their walking process (see Fig. 4). Then we use the MidHip point (i. e., feature point of the waist) to represent the walking position of participants in the image coordinate system.

Considering perspective, CIT is the time from pedestrians' MidHip keypoint at 350 pixels to 1000 pixels on the image's horizontal axis. CT is the time from 1000 pixels to 1330 pixels, and ACT is the time from 1330 pixels to 1600 pixels for the pedestrian's MidHip keypoint.

IV. RESULTS

A. Subjective Evaluations and Walking Times

Results of subjective evaluation factors (Q1 to Q6) and walking behavior factors (CIT, CT, and ACT) are presented by box plots in Fig. 5. In which, the red bars represent the median values.

Reviewing the eHMI conditions design (refer to Section III-E), our motivation is to collect various levels of psychological and behavioral data from pedestrians interacting with APMVs under the different eHMI conditions. To validate this motivation, the Friedman test was conducted to determine whether there are differences in each factor under different eHMI conditions since the Shapiro-Wilk test reported that the results for each factor did not conform to a normal distribution. In Table II, the results of the Friedman test show that there were significant differences in four eHMI conditions for all factors except for Q5 and Q6. Note that the primary focus of this paper is on the causal discovery of pedestrians' psychological states leading to walking behaviors in pedestrian-APMV interactions. This paper does not aim to investigate the impacts of different eHMI on pedestrian-APMV interactions. Thus, post-hoc comparisons of these

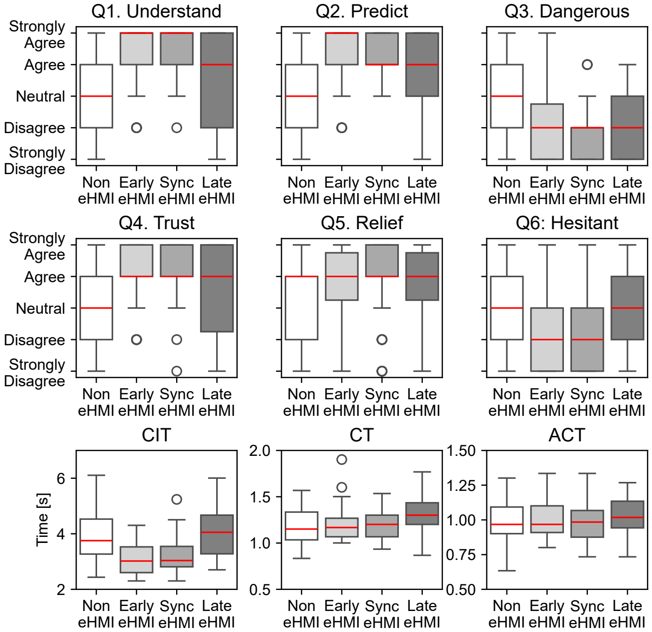

Fig. 5: Results of subjective evaluations (Q1 to Q6) and walking behavior factors (CIT, CT and ACT).

TABLE II. Friedman test results for each subjective evaluation factor and walking behavior factor. *:$p < .05$, **:$p < .01$, ***:$p < .001$.

	W	ddof1	ddof2	F	p
Q1	0.444	2.889	49.111	13.580	<.001 ***
Q2	0.385	2.889	49.111	10.630	<.001 ***
Q3	0.227	2.889	49.111	4.979	.005 **
Q4	0.322	2.889	49.111	8.082	<.001 ***
Q5	0.077	2.889	49.111	1.413	.251
Q6	0.134	2.889	49.111	2.629	.063
CIT	0.439	2.889	49.111	13.279	<.001 ***
CT	0.253	2.889	49.111	5.756	.002 **
ACT	0.159	2.889	49.111	3.206	.033 *

factors between eHMI conditions are not reported in this paper.

Furthermore, Table III presents descriptive statistics for the results related to subjective evaluation factors and walking behavior factors in all conditions of the eHMI. Throughout the experiment, a total of 216 datasets, each comprising nine factors, were collected. Assessing the maximum, minimum values, and standard deviation (std) of data for each factor, the collected dataset is deemed extensive, non-sparse, and conducive to causal discovery. The table also includes Pearson correlation coefficients between each pair of factors. This helps us to preliminarily explore potential relationships between factors in pairs.

B. Causal Discovery via DirectLiNGAM

For the data collected as described in the previous section, we employed DirectLiNGAM for causal discovery. Based on the hypothesized model represented in Fig. 1, we establish prior knowledge of DirectLiNGAM. Specifically, Q1 is designated as an exogenous variable, i. e., independent variable, indicating that it is considered a factor within the model not influenced by other variables but capable of influencing them. Simultaneously, the three factors of walking behavior, i. e., CIT, CT and ACT, have been designated as endogenous variables, i. e., dependent variable, indicating that it is considered a factor within the model influenced by other variables but not capable of influencing them.

The result of causal discovery through DirectLiNGAM is presented in Fig. 6. The blue node represents an exogenous variable and the red nodes represent endogenous variables. The arrows between nodes indicate the causal direction, and the numbers next to the edges represent the direct causal effects, i.e., the corresponding element of the adjacency matrix.

C. Statistical Reliability of DAG Based on the Bootstrap

The bootstrap method was used to further analyze the reliability of the causal discovery results [18]. In specific terms, 1000 bootstrap replications were performed to generate new datasets from the collected dataset, each maintaining the same data size. For 1000 generated new dataset, DirectLiNGAM was employed 1000 times to conduct causal discovery and estimate the adjacency matrix $\boldsymbol{A}$ separately. Subsequently, the 1000 adjacency matrices of $\boldsymbol{A}$ were aggregated to calculate the probability of occurrence for non-zero elements. This represents the reproducibility probability of the presence of direct effects indicating direct causal relationships between the corresponding pairs of factors. Figure 7 shows a DAG formed by directed edges with occurrence probabilities exceeding 50%, as determined through 1000 bootstrap iterations. The reproducibility probability for each directed edge is labeled next to it.

V. DISCUSSION

Figures 6 and 7 show the results of causal discovery. In specific, they showed a high reliable direct causal relationship from Q1 to Q2 (reproducibility probability = 100%), i.e., the understanding of APMV's driving intentions leaded to the prediction of APMV's driving behaviors This result aligned with the process of situation awareness [13] in the hypothesized model (see Fig.1).

Additionally, Q5 was confirmed to be directly generated from Q1 only (reproducibility probability = 72%) , i. e., the understanding of APMV's driving intentions directly leaded to the sense of relief experienced by pedestrians. This result was not align the inference discussed in [9], where the interaction between pedestrians and AV suggested that a sense of danger and trust in AV led to the inference of a sense of relief. The potential reason for this discrepancy could be attributed to the limited data collected in this study, leading to a sparse distribution of results for Q5 as shown in Table III.

Moreover, Fig. 6 shows that Q4 is directly generated by both Q1 and Q2, i. e., participants' trust in APMV tends to increase if the driving intentions of APMV could be easily understood and the driving behavior could be easily predicted. As shown in Fig. 7, the results of 1000 Bootstrap iterations indicate a 100% of reproducibility probability for the direct edges from Q1 to Q4 but the reproducibility probability of the directed edges from Q2 to Q4 was lower than 50%.

TABLE III. Descriptive Statistics and Pearson correlation coefficients of factors under all conditions. For Q1 to Q6, 1="strongly disagree", 2="disagree", 3="neutral", 4="agree", and 5="strongly agree". Unit of CIT, CT and ACT is second.

	Descriptive Statistics					Pearson Correlation Coefficient								
	N	mean	std	min	max	Q1	Q2	Q3	Q4	Q5	Q6	CIT	CT	ACT
Q1	216	3.84	1.26	1.00	5.00	1.00	0.85	-0.73	0.78	0.32	-0.70	-0.28	-0.03	-0.03
Q2	216	3.84	1.15	1.00	5.00	0.85	1.00	-0.76	0.75	0.30	-0.73	-0.34	-0.06	-0.10
Q3	216	2.22	1.20	1.00	5.00	-0.73	-0.76	1.00	-0.79	-0.28	0.77	0.18	-0.06	-0.07
Q4	216	3.70	1.20	1.00	5.00	0.78	0.75	-0.79	1.00	0.30	-0.75	-0.20	0.00	0.05
Q5	216	3.65	1.21	1.00	5.00	0.32	0.30	-0.28	0.30	1.00	-0.24	-0.10	0.03	0.10
Q6	216	2.60	1.31	1.00	5.00	-0.70	-0.73	0.77	-0.75	-0.24	1.00	0.33	0.01	-0.03
CIT	216	3.58	0.85	2.30	6.11	-0.29	-0.35	0.19	-0.20	-0.11	0.34	1.00	0.30	0.22
CT	216	1.23	0.21	0.83	2.37	0.00	-0.03	-0.06	0.03	0.05	-0.06	0.30	1.00	0.73
ACT	216	1.00	0.13	0.63	1.33	-0.04	-0.11	-0.05	0.04	0.10	0.00	0.22	0.73	1.00

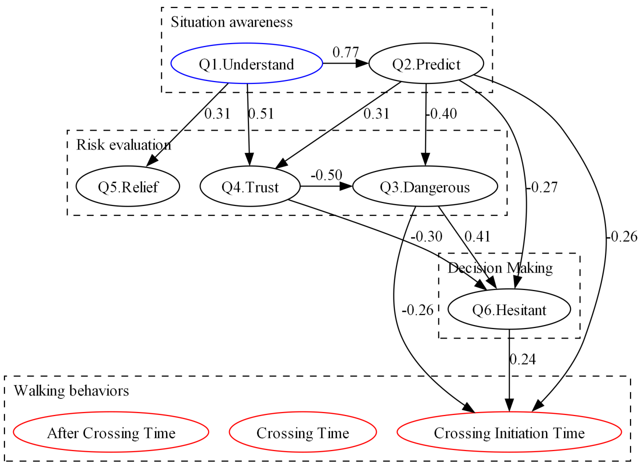

Fig. 6: DAG estimated by DirectLiNGAM. Blue node represents the designated exogenous variable, while red node represents the designated endogenous variable.

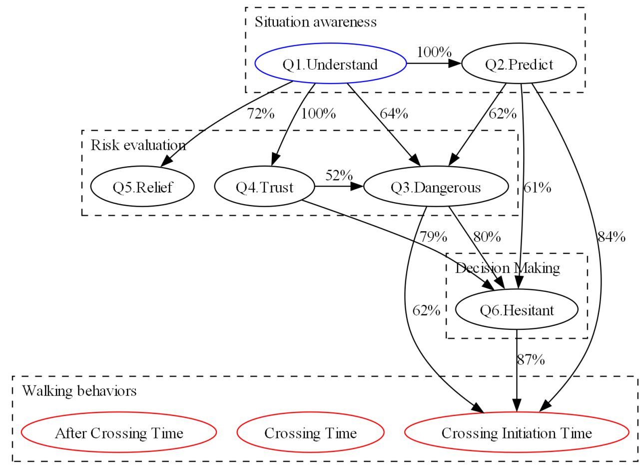

Fig. 7: Directed edges with a reproducibility probability exceeding 50% through 1000 bootstrap iterations. Blue node represents the designated exogenous variable, while red node represents the designated endogenous variable.

This indicates that pedestrians' understanding of the driving intentions of APMV was a significant factor in their trust in APMV. This is consistent with the results of [19] and [9], although it discussed the driver–AV interactions. This result also suggests that for pedestrians to trust APMV appropriately, avoiding both overtrust and undertrust, it is crucial to help pedestrians establish a correct mental model of APMV to assist them in achieving accurate situation awareness.

Next, Q3 was directly generated by both Q2 and Q4 shown Fig. 6, meaning that when predicting the driving behavior of APMV is challenging or when trust in APMV is low, participants may perceive danger. This result aligns with the discussion in [9]. Additionally, Fig. 7 further reveals that, through bootstrap resampling, high reproducibility probabilities were discovered not only for the directed edges from Q2 and Q4 to Q3 but also for the directed edge from Q1 to Q3. Similar to this result, Liu et al. calculated the conditional probability of a sense of danger given the understanding of APMV's driving intentions and found a causal relationship between them [20], without analyzing the prediction aspect. De Clercq et al. also reports that easily understandable driving intentions can alleviate pedestrians' sense of danger [8]. Based on the results of this study, the understanding of APMV's driving intentions indirectly leads to the sense of danger through the prediction of APMV's driving behavior.

Figures 6 and 7 show that Q6 was generated from Q2 (reproducibility probabilities = 61%), Q3 (reproducibility probabilities = 80%) and Q4 (reproducibility probabilities = 79%), i. e., the participants were less to hesitate in decision-making when they find it easier to predict the driving behavior of APMV, have higher trust in APMV, and perceive less danger associated with it.

Surprisingly, in pedestrians' walking behaviors, except for CIT, DirectLiNGAM inferred that ACT and CT are independent variables with no causal relationships with other factors. This might be due to the narrow width of the road that pedestrians need to cross and the short distance to the end of the road after crossing (see Fig. 4).

On the other hand, Figs. 6 and 7 show that Q2 (reproducibility probabilities = 84%), Q3 (reproducibility probabilities = 62%) and Q6 (reproducibility probabilities = 62%) directly lead to CIT. Specifically, when the behavior of APMV is challenging to predict, when there is no perceived danger, and when hesitation is felt, it results in an increase in CIT. The findings of a similar study suggest a shorter CIT for pedestrians when they can easily comprehend the autonomous vehicle's intentions through the eHMI [10]. In our research, a more easily understood driving intent indirectly contributes to a reduced CIT by facilitating the prediction of driving behavior.

Additionally, there seems to be a perplexing result where the direct causal effect from Q3 to CIT was negative, i. e., -0.26, shown in Fig. 6, but the correlation coefficient between Q3 and CIT was positive, i. e., 0.19, shown in Table III. This could be considered that the correlation analysis focuses on the relationship between two variables, whereas DirectLiNGAM based on SEM involves a multivariate linear regression analysis.

Overall, the general flow of the results from causal discovery aligns with the hypothesized model (see Fig. 1), where the outcomes of situation awareness lead to risk evaluation, the outcomes of risk evaluation lead to hesitation in decision-making, and the outcomes of hesitation lead to walking behaviors.

A. Limitations

While this study employed the bootstrap method to artificially augment the dataset for assessing the reliability of causal discovery, it's important to note that the collected dataset from 18 participants. Therefore, the limited size of the dataset may impose constraints on the results of causal discovery. Furthermore, since the participants were all in 20s and 30s, the results may not necessarily be applicable to children and elderly individuals.

VI. Conclusion

This study aims to investigate the causal relationships from pedestrians' psychological states to their walking behavior during interactions with APMV. A subjective experiment was conducted, where various eHMIs were designed to induce participants to experience different levels of subjective feelings and generate corresponding walking behaviors. By employing DirectLiNGAM to analyze the collected data for causal discovery, the results of causal discovery align with the hypothesized model depicted in Fig. 1. Furthermore, the experimental results have enriched the detailed causal relationships within the hypothesized model, i. e., the outcomes of situation awareness lead to the sense of dangerous, trust in APMV and sense of relief; the outcomes of situation awareness, the sense of danger and trust in APMV lead to hesitation in decision-making; and the outcomes of situation awareness, the sense of danger and hesitation lead to walking behaviors.

In future work, we plan to increase the number of participants to obtain more reliable causal discovery results. Based on this study, we will also further exploring design guidelines for eHMI and strategies for calibrating people's trust in APMV.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Numbers 20K19846, 22H00246 and 24H00298, Japan.

References

[1] SAE Technical Standards Board, "J3016b: Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles." SAE International, 2018, pp. 1–35.

[2] H. Liu, Y. Li, Z. Zeng, H. Cheng, C. Peng, and T. Wada, "Is silent ehmi enough? a passenger-centric study on effective ehmi for autonomous personal mobility vehicles in the field," *International Journal of Human–Computer Interaction*, 2024.

[3] L. Y. Morales, T. Miyashita, and N. Hagita, "Social robotic wheelchair centered on passenger and pedestrian comfort," *Robotics and Autonomous Systems*, vol. 87, pp. 355–362, 2017.

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

[5] H. Liu, T. Hirayama, L. Y. M. Saiki, and H. Murase, "Implicit interaction with an autonomous personal mobility vehicle: Relations of pedestrians' gaze behavior with situation awareness and perceived risks," *International Journal of Human–Computer Interaction*, vol. 39, no. 10, pp. 2016–2032, 2023.

[6] H. Liu, T. Hirayama, L. Y. M. Saiki, and H. Murase, "What timing for an automated vehicle to make pedestrians understand its driving intentions for improving their perception of safety?" in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 2020, pp. 462–467.

[7] C.-M. Chang, K. Toda, D. Sakamoto, and T. Igarashi, "Eyes on a car: an interface design for communication between an autonomous car and a pedestrian," in *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications*, 2017, pp. 65–73.

[8] K. de Clercq, A. Dietrich, J. P. N. Velasco, J. de Winter, and R. Happee, "External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions," *Human Factors*, vol. 61, no. 8, pp. 1353–1370, 2019.

[9] H. Liu, T. Hirayama, and M. Watanabe, "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," in *2021 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2021, pp. 748–754.

[10] A. Loew, J. Graefe, L. Heil, A. Guthardt, A. Boos, A. Dietrich, and K. Bengler, "Go ahead, please!—evaluation of external human—machine interfaces in a real-world crossing scenario," *Frontiers in Computer Science*, vol. 4, p. 863072, 2022.

[11] Y. M. Lee, R. Madigan, C. Uzondu, J. Garcia, R. Romano, G. Markkula, and N. Merat, "Learning to interpret novel ehmi: The effect of vehicle kinematics and ehmi familiarity on pedestrian'crossing behavior," *Journal of safety research*, vol. 80, pp. 270–280, 2022.

[12] D. Dey, A. Matviienko, M. Berger, B. Pfleging, M. Martens, and J. Terken, "Communicating the intention of an automated vehicle to pedestrians: The contributions of ehmi and vehicle behavior," *it-Information Technology*, vol. 63, no. 2, pp. 123–141, 2021.

[13] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human factors*, vol. 37, no. 1, pp. 32–64, 1995.

[14] G. J. Wilde, "The theory of risk homeostasis: implications for safety and health," *Risk analysis*, vol. 2, no. 4, pp. 209–225, 1982.

[15] S. Shimizu, P. O. Hoyer, A. Hyvärinen, A. Kerminen, and M. Jordan, "A linear non-gaussian acyclic model for causal discovery." *Journal of Machine Learning Research*, vol. 7, no. 10, 2006.

[16] S. Shimizu, T. Inazumi, Y. Sogawa, A. Hyvarinen, Y. Kawahara, T. Washio, P. O. Hoyer, K. Bollen, and P. Hoyer, "Directlingam: A direct method for learning a linear non-gaussian structural equation model," *Journal of Machine Learning Research-JMLR*, vol. 12, no. Apr, pp. 1225–1248, 2011.

[17] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh, "Openpose: Realtime multi-person 2d pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[18] K. Thamvitayakul, S. Shimizu, T. Ueno, T. Washio, and T. Tashiro, "Bootstrap confidence intervals in directlingam," in *2012 IEEE 12th International Conference on Data Mining Workshops*. IEEE, 2012, pp. 659–668.

[19] L. Petersen, L. Robert, X. J. Yang, and D. M. Tilbury, "Situational awareness, drivers trust in automated driving systems and secondary task performance," *arXiv preprint arXiv:1903.05251*, 2019.

[20] H. Liu, T. Hirayama, L. Y. M. Saiki, and H. Murase, "Implicit interaction with an autonomous personal mobility vehicle: Relations of pedestrians' gaze behavior with situation awareness and perceived risks," *International Journal of Human–Computer Interaction*, vol. 39, no. 10, pp. 2016–2032, 2023.