Article:
Wang, Y., Srinivasan, A.R., Lee, Y.M. orcid.org/0000-0003-3601-4191 et al. (1 more author) (2025) Modeling Pedestrian Crossing Behavior: A Reinforcement Learning Approach With Sensory Motor Constraints. IEEE Transactions on Intelligent Transportation Systems. ISSN: 1524-9050

https://doi.org/10.1109/tits.2025.3581693

Modeling Pedestrian Crossing Behavior: A Reinforcement Learning Approach with Sensory Motor Constraints

Yueyang Wang, Aravinda Ramakrishnan Srinivasan, Yee Mun Lee, and Gustav Markkula

***Abstract*—Understanding pedestrian behavior is crucial for the safe deployment of Autonomous Vehicles (AVs) in urban environments. Traditional pedestrian behavior models often fall into two categories: mechanistic models, which do not generalize well to complex environments, and machine-learned models, which generally overlook sensory-motor constraints influencing human behavior and which are thus prone to fail in unseen scenarios. We hypothesize that sensory-motor constraints, fundamental to how humans perceive and interact with their surroundings, are essential for realistic simulations. Thus, we introduce a constrained reinforcement learning (RL) model that simulates the crossing decision and locomotion of pedestrians. Our model includes human sensory constraints, giving the agent imperfect information about the environment, and human motor constraints incorporated through a bio-mechanical model of walking. We gathered data from a human-in-the-loop experiment to understand pedestrian behavior. The findings reveal several behavioral patterns not addressed by existing pedestrian models, regarding how pedestrians adapt their walking speed to the kinematics and behavior of the approaching vehicle. Our model successfully captures these human-like walking speed patterns, enabling us to understand these patterns as a trade-off between time pressure and walking effort. Importantly, the model with both sensory and motor constraints performed better than models only incorporating one of the two. Additionally, behavioral patterns related to external human-machine interfaces and light conditions were also captured by the model. Overall, our results not only demonstrate the potential of constrained RL in modeling pedestrian behaviors but also highlight the importance of sensory-motor mechanisms in modeling pedestrian-vehicle interactions.**

***Index Terms*—Reinforcement learning, Noisy perception, Road user interaction, Pedestrian behavior, Sensory-motor constraints.**

I. Introduction

AUTONOMOUS vehicles (AVs) have attracted considerable attention from the public and play a crucial role in developing the intelligent transportation system of tomorrow. A critical aspect of integrating AVs into the urban environment is their ability to interact safely and smoothly with other road users [1]. Among these road users, pedestrians exhibit complex and often unpredictable behaviors. Thus, to improve the interaction between AVs and pedestrians, researchers have tried different approaches. For example, external Human-Machine Interfaces (eHMIs) have been proposed as a means to communicate AV intentions [2, 3], and their effectiveness has been proven [4, 5]. Other researchers have tried to develop models that can help us understand and predict pedestrian behavior [6]. These models employs primarily two approaches: mechanistic models and machine learning (ML) models. Mechanistic models, including cognitive models grounded in psychology and neuroscience, provide detailed insights into the cognitive processes underlying pedestrian behaviors but often struggle with the complexity and diversity of real-world scenarios [7, 8]. ML models, in contrast, use data-driven techniques to learn from large datasets and predict pedestrian movements effectively. However, they require extensive labeled data, lack interpretability, and sometimes do not generalize well outside their training datasets [9, 10]. In addition, ML models primarily aim to improve motion prediction accuracy using metrics such as Root Mean Squared Error (RMSE) and Average Displacement Error (ADE). However, accurate motion prediction alone is not sufficient to ensure optimal AV performance. AVs must also consider social interaction and understand the decision-making processes of other road users. By doing so, AVs can exhibit more adaptive and socially-aware driving behaviors, leading to safer and more efficient interactions in complex traffic environments [11].

Our previous work has focused on addressing these challenges by developing a pedestrian model that integrates the strengths of both cognitive and ML approaches [12]. Specifically, we have explored the use of reinforcement learning (RL) to model the binary decision of go/no-go in pedestrian crossing scenarios when interacting with an approaching vehicle, incorporating theory about visual perception to capture realistic human-like road crossing decisions.

In the present work, we aim to further this approach by considering additional aspects of visual perception (eHMI and lighting conditions), and crucially extending it with a biomechanical model of walking, allowing us to address also motor aspects of pedestrian behavior. By integrating these sensory-motor mechanisms into RL models, we attempt to capture a range of behavioral patterns observed in a controlled

Corresponding author: Yueyang Wang.

This work was supported by U.K. Engineering and Physical Sciences Research Council under Grant EP/S005056/1.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of the University of Leeds under Application No. 0536.

Y. Wang is with the Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT UK (E-mail: mn20yw2@leeds.ac.uk).

A. Srinivasan was with the Institute for Transport Studies, University of Leeds, LS2 9JT Leeds, U.K. He is now with Chair and Institute of Highway Engineering (ISAC), RWTH Aachen University, Aachen, 52074, Germany (E-mail: aravinda@isac.rwth-aachen.de).

Y. Lee and G. Markkula are with the Institute for Transport Studies, University of Leeds, LS2 9JT Leeds, U.K (E-mail: Y.M.Lee@leeds.ac.uk; G.Markkula@leeds.ac.uk).

experiment on pedestrian road-crossing.

II. Background

A. Pedestrian crossing behaviors

1) Observed behavioral patterns in pedestrian crossing: In this paper, we define a behavioral pattern as a relationship between a dependent variable that reflects pedestrian behavior and an independent variable influencing it. Below, we introduce several key behavioral patterns related to pedestrian crossing as reported by previous literature. Some of these behavioral patterns were modeled in our previous study [12], while others remain less explored. One behavioral patterns we modeled is the gap acceptance rate, which refers to the rate at which pedestrians accept the time or distance gap between themselves and approaching vehicles. Another is the crossing initiation time (CIT), defined as the time between the gap appearing and the pedestrian starting to cross. Both gap acceptance and CIT increase with the time to arrival (TTA) of the vehicle, as well as with the vehicle's speed [13]. Another important behavioral patterns regarding the pedestrian crossing speed, was not captured by our previous crossing decision model [12]. This metric tends to decrease as the time gap increases, suggesting a compensatory behavior where pedestrians walk faster when accepting shorter gaps [14].

2) Sensory-motor mechanisms in pedestrian crossing: Pedestrian crossing decisions are influenced by a variety of sensory-motor mechanisms, which impact how pedestrians perceive and interact with their environment. In this study, we focused on noisy perception, looming aversion, time pressure, walking effort, and ballistic speed control.

a) Noisy perception: Human perception of the world is inherently noisy and imperfect [15]. This noisiness can be due to several factors, such as individual differences in sensory acuity [16] and environmental conditions (e.g., poor lighting, weather). It has been argued that this visual limitation affected the pedestrian crossing decision [17]. For example, noisy perception can lead to errors in estimating the speed and distance of oncoming vehicles, making it challenging for pedestrians to accurately judge safe crossing opportunities.

b) Looming aversion: Visual looming describes the perceived growth of an object's size as it approaches [18]. The aversion to looming refers to the tendency of individuals to react more strongly to objects that appear to be rapidly approaching [19]. This phenomenon is rooted in the perceptual system's sensitivity to motion cues that signal potential threats, and influences pedestrian crossing decision. For example, [13] explained speed-dependent gap acceptance behavior by using looming aversion.

c) Time pressure: Time pressure has a significant effect on pedestrian crossing behavior [13]. For example, when pedestrians are under time constraints, either due to the rapid approach of a vehicle or a short signal phase [14], they tend to initiate crossing faster and adopt higher walking speeds.

d) Walking effort: Individuals choose their walking speed accounting for not only the time but also the energy spent on walking [20]. Researchers have considered the walking effort in gait modeling works [21, 22]. However, the impact of walking effort has not been extensively investigated in the existing pedestrian modeling literature. Considering the trade-off between the energy and time costs can help to better predict the walking dynamics. For example, studies have shown that older pedestrians often exhibit more conservative crossing behaviors, partly due to the increased walking effort required and the need to ensure their safety [23].

e) Ballistic speed control: Early researchers noted that humans tend to adjust their sensorimotor movements through intermittent, ballistic control, rather than continuous adjustments [24, 25]. Here, we adopt this ballistic perspective in the context of walking, assuming that each walking step is ballistic. In reality, humans can to some extent alter their walking trajectory in the middle of a walking step [26], but in the present model we neglect this behavior.

B. Pedestrian crossing models

The development of models to predict and understand pedestrian crossing behavior has been a significant focus of research for many years. These models vary in complexity and approaches, capturing different aspects of pedestrian decision-making processes and walking dynamics.

1) Logistic regression models: Logistic regression models, including more recent cognitive cue-based models like the one proposed in [13], aim to predict whether a pedestrian will cross the street under certain conditions. These models typically consider factors such as vehicle speed, and distance [27]. However, these models primarily focus on crossing decisions without considering the subsequent walking dynamics, making them difficult to integrate into more comprehensive simulations of traffic and pedestrian behavior.

2) Mechanistic models: More complex mechanistic models, such as evidence accumulation models, simulate the human decision-making process by considering how pedestrians gather information over time [28, 29]. These models attempt to capture the gradual accumulation of sensory and contextual data that leads to crossing decisions. They also provide a deeper understanding of the cognitive processes involved, and in some cases also include some limited consideration of walking dynamics [8]. However, integrating multiple cognitive theories into a unified model is challenging because these theories often have differing assumptions, involve complex interactions, require scalable and adaptable frameworks, and increase model complexity, which can reduce practical usability. This complexity makes it difficult to capture a broader range of behaviors and apply such models to more complex situations and environments [8].

3) Machine Learning Models: ML models leverage large datasets and advanced algorithms to predict pedestrian behavior with high accuracy and have shown good performance in real-time pedestrian trajectory prediction [10]. While the accuracy of these ML models can be impressive, they share several limitations. First, they often act as 'black boxes', meaning they may fail to reveal the reasons or mechanisms behind those behaviors [30]. In addition, the 'black box' nature of ML models in autonomous driving poses risks, as their lack of transparency can lead to unexplainable mistakes [31].

Second, high-level statistical accuracy does not guarantee that the models capture those behaviors that matter to humans [32]. Third, the absence of a theoretical grounding of the behavior can sometimes lead to performance issues. While they perform well on the training data, they may struggle to generalise to new and unseen data [33]. In addition, ML models heavily rely on large training datasets. Whereas collecting such extensive data under all possible road conditions is almost an impossible task, leading to critical scenarios being underrepresented or missing in most datasets [34].

4) Models based on (bounded) optimality: Another approach to modeling human behavior is *bounded optimality* [8, 12, 35, 36], based on the assumption that humans behave optimally with respect to a utility or cost function, but with constraints imposed by the human cognition and body [37, 38]. By using RL as a method for solving the bounded optimality problem, this approach integrates the strengths of both cognitive models and ML models.

Different from traditional data-driven ML algorithms, which typically learn directly from large datasets without iterative interaction with the environment, RL offers a paradigm wherein an agent interacts with a dynamic environment, and the optimal policy will be derived through trial-and-error [39]. This approach is particularly suited for modeling tasks that involve sequential decision-making, such as pedestrian crossing. Crossing decisions require a series of interdependent sensory and motor actions, as pedestrians continuously assess environmental factors like vehicle speed and proximity, while adjusting their own movements accordingly. By incorporating models of human sensory and motor mechanisms, RL can be used to learn bounded optimal behavior under these constraints.

Our previous work using this method captured the pedestrian crossing decisions when interacting with one vehicle, and captured several observed behavioral patterns [12]. However, research gaps remain in the previous model. Key questions include whether our integrated cognitive and RL approach can generalise to a wider range of situations and behavioral patterns, and whether the model can be expanded to include motor constraints, which would allow us to more accurately represent the physical execution of walking actions during road crossing.

III. Methods

A. Experiment

To develop and test our model, we used a dataset from an experiment conducted in the University of Leeds Highly Immersive Kinematic Experimental Research (HIKER) laboratory.

The experiment is described in full in [40], below we provide a summary for completeness. In this study, the task of the participant was to cross the road between two approaching vehicles safely – as shown in panel (b) of Fig. 1. A mixed design was used, with five within-participant variables: (i) the initial speed v_0 of the approaching vehicles (25/30 mph); (ii) the time gap t_0 between the vehicles (3/5 s); (iii) the yielding behavior of the second vehicle with a constant deceleration rate a (yielding/non-yielding); (iv) the presence of eHMI while yielding (present/absent); and (v) the time of day (daytime/nighttime); and one between-participant variable: participants' age (younger/older) [40]. In this study, the eHMI was presented as a cyan color light around the windscreen in some of the yielding trials. Participants were informed that the presence of eHMI means that the approaching vehicle was signaling *'I am yielding'*. Prior to the experiment, all participants signed the consent form agreeing to take part in the study. Ethical approval was obtained from the University of Leeds Research Committee.

The recorded data used from the experiment include participants' crossing decisions, CIT, and walking speeds; these quantities were all estimated from a body tracker located on the participant's head, measuring XYZ position at a sampling rate of 50 Hz. Furthermore, we classified crossing behaviors in yielding scenarios as either early or late. This classification relies on the concept of bimodal crossing, where pedestrian responses can typically be categorized into two distinct modes based on their timing relative to vehicle behaviors, with pedestrians choosing to cross either well before a vehicle slows down or waiting until it is clearly safe to do so [28]. In our study, early and late crossings are defined based on the CIT of the pedestrian crossing. Specifically, early crossings occur when crossings are initiated at the first peak of the bimodal CIT distribution, while late crossings correspond to those initiated at the second peak. The classification threshold was determined based on vehicle deceleration: in yielding scenarios, this point corresponds to the moment vehicle speed drops to two-thirds of the initial speed. This threshold approximately aligns with the lowest point between the two peaks. For a more detailed explanation of this method, please refer to [40]. It should be noted that the distinction between early and late crossings is applicable only in yielding scenarios.

B. Model

1) Sensory-motor mechanisms: In this paper, we explore the influence of four sensory-motor mechanisms on the crossing behavior as detailed in Section II-A2, noisy perception, visual looming, time pressure, walking effort and ballistic speed control; as illustrated in Fig. 1) (a).

a) Noisy perception: The agent's perception of environmental distances includes a constant Gaussian angular noise, σ_{v}, which affects the estimation of vehicle positions and velocities. A higher σ_{v} means more noise in the agent's perception system. This model incorporates a Kalman filter to adaptively refine these estimates, leveraging Bayesian methods

TABLE I
VEHICLE APPROACH SCENARIOS IN THE EXPERIMENT.

Scenario type	v_0 (mph)	τ_0 (s)	a (m/s^2)
Constant speed	25	3	N/A
	25	5	N/A
	30	3	N/A
	30	5	N/A
Yielding	25	3	-2.3
	25	5	-2.3
	30	3	-2.3
	30	5	-2.3

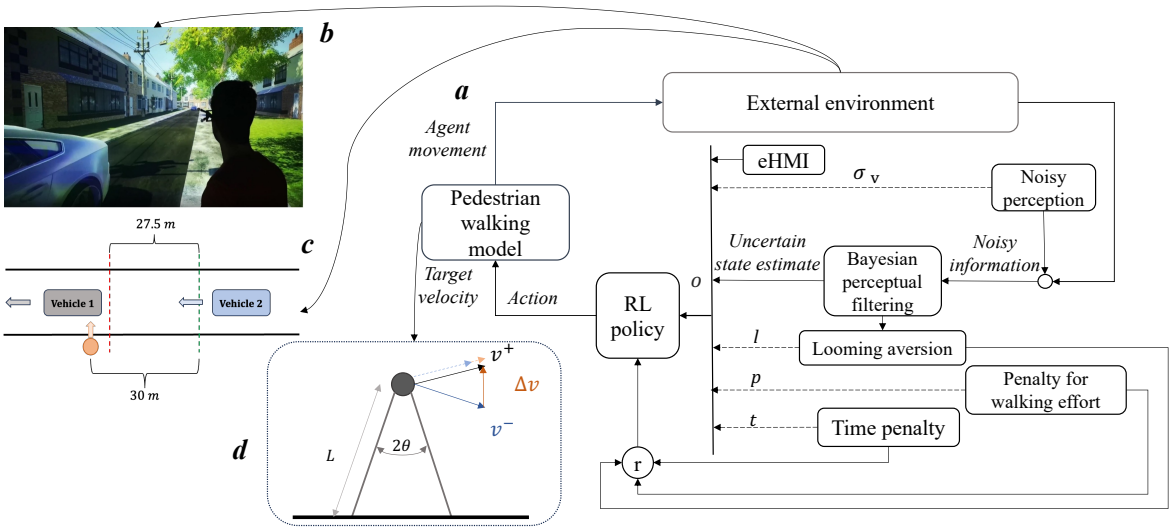


Fig. 1. Panel (a) is the Model framework. Panel (b) is the virtual environment. Panel (c) is a schematic of the deceleration procedure used in this study. Panel (d) is the walking model, adopted from [20].

to approximate the state of the environment refer to the previous paper for mathematical details [12].

b) Visual looming: This phenomenon is mathematically represented as inverse τ—the ratio of a vehicle's optical expansion rate to its size on the observer's retina, which is an estimate of the inverse TTA [41, 42]. We defined looming aversion as follows: $L = c \cdot \frac{1}{\hat{\tau}}$, where c is the weight of the looming aversion, if $v > 0$; otherwise 0, and $\hat{t}$ Kalman-filter estimated TTA of the vehicle. The higher c means the agent has a stronger preference for avoiding visual looming.

c) Walking effort: To account for walking effort, we adopt the biomechanical model of walking from [20]. The equation for the new velocity v_i^+ is:

$$v_i^+ = v_i^- \cos 2\theta + \sqrt{2u_i} \sin 2\theta \tag{1}$$

where v_i^- is the initial walking speed, v_i^+ is the new walking speed after the action, and 2θ is the angle between two legs. The term $\sqrt{2u_i}$ represents the velocity component due to the exerted effort. The derivation of Equation 1 starts with the assumption that the pedestrian's leg motion can be described as a pendulum (see panel (d) of Fig. 1), and that the speed change is caused by the effort u_i. This effort required for this change is related to the kinetic energy:

$$u_i = \frac{1}{2} m (\Delta v)^2 \tag{2}$$

and solving for Δv: $\Delta v = \sqrt{2u_i/m}$. Since u_i is effort per unit mass, this simplifies to: $\Delta v = \sqrt{2u_i}$. Therefore, the effort u_i required for this change is given by:

$$u_i = \frac{\left(v^- \cos(2\theta) - v^+\right)^2}{2 \cdot \sin^2(2\theta)} \tag{3}$$

The total walking effort, E_w, is then calculated as: $E = \beta \cdot u_i$, where β is a parameter that scales the walking effort for different individuals. A higher β means the agent has a stronger preference for saving energy during crossing.

d) Ballistic speed control: As mentioned in Section II-A2, we assumed that the agent adjust their movement using the ballistic speed control. Therefore, upon selecting an action, the agent employs ballistic speed control, which means the agent maintains the acceleration rate a needed to change its speed to the desired value, which is defined as $a = \frac{V_t - V_{t-1}}{t}$, where V_t and V_{t-1} represent the velocities of the agent at two consecutive decision points. The time interval T_{step} is calculated based on the relation between step length, velocity, and frequency, with the preferred speed v and step length s following the equation $s = {v_{\text{ped}}}^{0.42}$ [43]. Consequently, the duration of a walking step is given b $T_{\text{step}} = {v_{\text{ped}}}^{-0.58}$.

2) Reinforcement learning problem: Our RL environment is an example of a Partially Observable Markov Decision Process (POMDP), where the agent does not have direct access to the true state; rather, it receives observations that may only partially or noisily reflect the actual state. The POMDP is represented by a tuple $< S, A, T, R, O >$, where S is a set of states, A is a set of actions, T is the transition function, R is the reward function and O is the set of possible observations received by the agent.

a) State: At each time step t, the environment is in a state $s_t \in S$, which includes the position and velocity of both the pedestrian (agent) and two vehicles. All model variants in our study share the same state space: the pedestrian's position $x_{\text{ped}}, y_{\text{ped}}$, the pedestrian's velocity v_{ped}, the vehicles' positions $x_{\text{veh}_1}, x_{\text{veh}_2}, y_{\text{veh}_1}, y_{\text{veh}_2}$, the vehicles' velocities v_1, v_2, and time step t. The simulation updates the state every 0.1 s, a step size chosen as a tradeoff between the computational cost and the accuracy to represent dynamic interactions between vehicles and pedestrians.

b) Action: At the completion of each walking step T_{step}, the agent executes an action a_t from set A. In this study, the action space in this study is $A = \{0.1, 0.2, \ldots, 2\}$ m/s. Each action value corresponds to a different desired walking speed, enabling the agent to adjust its velocity. Upon selecting an action, the agent employs ballistic speed control, which means the agent maintains the acceleration rate needed to change its speed to the desired value, which is defined as $a = \frac{V_t - V_{t-1}}{t}$. After the decision is made, the simulation progresses until the execution of the walking step is completed, at which point the agent can make the next decision.

c) Transition: The transition function defines how the current state s_t changes to the next state s_{t+1} based on the

action a_t. As the decision-making of the pedestrian is the main focus of this study, in line with the experimental setup, the vehicle in our simulation maintains a predefined behavior pattern and does not dynamically respond to the pedestrian's actions. This approach allows us to isolate and analyze the pedestrian's decision-making process. The vehicle's movement follows kinematic equations with the speeds and accelerations corresponding to the scenario in question. The pedestrian's walking speed is updated with the ballistic walking step acceleration as explained above. The simulation ends when a collision occurs or the agent crosses the road safely.

d) Reward: According to the experiment, we assumed that the agent wanted to cross safely, while minimizing time losses, energy losses, and any discomfort from experiencing high levels of visual looming.

The reward function r was designed accordingly: $r = A - C - E - L$, where $A = 20 - T$ if the agent successfully crosses the road, otherwise $A = 0$, and where T is the weighted time penalty, $T = \alpha \cdot t$, where t is the time elapsed in the episode and α is the scaling factor of the time pressure; with higher α making the agent more likely to cross the road quickly; C is 20 if collision, otherwise 0; $L = c \cdot \frac{1}{\tau}$ if $v > 0$, otherwise 0, and $E = \beta \cdot u_i$, where E is the penalty for walking effort. The reward r is bounded within the range $[-20, +20]$ to prevent extreme values from affecting the model training. The value of 20 was initially determined through manual testing of the model in a simplified scenario without sensory-motor constraints, where the agent attempted to cross the road as quickly as possible without colliding. This testing helped establish a baseline for the reward scale, ensuring that the reward values yielded qualitatively human-like pedestrian behavior.

e) Observation: As previously mentioned, the agent observes their own state and Kalman estimates of the vehicle's state. Moreover, because we have trials with eHMI on the vehicle, we also gave the input representing eHMI in the model, with 0 for eHMI off and 1 for eHMI on. In addition, since we have different parameters that influence the sensory-motor process, i.e., σ_v, α, β and c, we also gave these parameters as input to the RL policy, i.e., we are conditioning the RL on these parameters [44, 45], which we refer to as non-policy parameters to distinguish them from the parameters of the policy neural network (connection weights and biases).

3) Different Model Variants: To explore how the various mechanistic assumptions in the model affect the generated pedestrian behavior, we tested different model variants:

SM: A model with all of the mechanistic assumptions described above, including both sensory assumptions (visual limitations and looming aversion) and motor assumptions (walking effort and ballistic speed control)

S: A model with the sensory assumptions but not the motor assumptions. When excluding the assumption about ballistic speed control, a more conventional RL approach of directly controlling the speed at each time step [46, 47, 48] was used.

M: A model with the motor assumptions but not the sensory assumptions.

4) Reinforcement learning: Proximal Policy Optimization (PPO) is a policy gradient RL algorithm that improves training stability and reliability by using a clipped objective function [49]. This method strikes a balance between exploration and exploitation, ensuring that updates to the policy are not too large. PPO is computationally efficient and straightforward to implement, making it a popular choice for various reinforcement learning tasks.

TABLE II
PPO MODEL CONFIGURATION

Parameter	Description
Policy Network	Multi-layer Perceptron (MLP)
Input Dimensions	Varies by model variants
Learning Rate	3×10^{-4}
Batch Size	64
Discount Factor	0.99
Network Structure	Two hidden layers with 128 and 64 neurons ReLU activation functions.

It should be noted that, in our study, PPO is used to learn an optimal behavioral policy for the simulated pedestrian, independent of the simulator study. The simulator study provides data on human road-crossing behavioral patterns, while PPO is applied separately to learn optimal behavior under different sensory and motor constraints.

For our implementation, we used the Stable Baselines 3 (SB3) library [50]. The model was trained for 3 million environment time steps using the PPO algorithm. The parameters for training are shown in Table II. The reward and total loss during training are shown in Fig. 2. Due to differences in POMDP structure between models, their final reward and loss values converge to different levels. As mentioned, we gave non-policy parameters as additional input to the RL network; for each new RL episode we sampled uniformly from these ranges: [σ_v (0-10), α (0-4), β (0-10), and c (0-10)]. In practice, this means that the RL is learning how the optimal policy varies across this space of non-policy parameters.

To ensure the agent did not learn trivial policies based on limited experimental conditions, we trained the RL agent with a wider range of kinematic conditions than the scenarios in the experiment. The initial speed was sampled from a uniform distribution between 8 m/s and 17 m/s, and the time gap was sampled from a uniform distribution between 0.1 s and 10 s.

5) Fitting non-policy parameters: As previously mentioned, we defined the magnitude of noise in the perception system, along with the weights for looming aversion, time pressure, and walking effort constraints, as non-policy parameters. A priori, we do not know the correct values for the non-policy parameters, which may also vary between participants in the experiment. Additionally, considering the day and night scenarios in the experiment, we hypothesized that visual noise differs between these scenarios. Consequently, for each participant, we wished to fit two σ_v values (one for day and

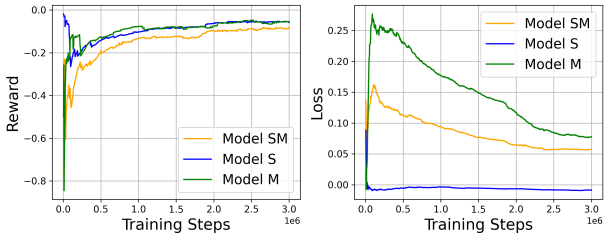


Fig. 2. Reward and training loss of different model variants.

Behaviours / Independent variable	Gap acceptance	Early/late crossing	CIT			Walking speed
			Non-yielding	Early crossing	Late crossing	
Time gap	S M -L -NP	S M -W -NP	S -L -W -B	S -B -NP	-L -NP	M -B -NP
Speed	-W -NP	-W -NP	-W -B	-L -W -B	M -W	
Day-time				-B		
eHMI					-NP	-B
Early crossing						M -L -W -NP

Fig. 3. Key behaviors in the experiment. Dark gray cells indicate statistically significant effects of the independent variable that were both observed in the experiment and captured by the SM model. Light gray cells indicate effects observed experimentally but not captured by the SM model. The symbols within the gray cells represent model variants that successfully reproduced the corresponding behavioral pattern: S (model with noisy perception and looming aversion), M (model with walking effort and ballistic control), –L (model without looming aversion), –W (model without walking effort), –B (model without ballistic control), and –NP (model without noisy perception).

one for night) and one value for each of the other three non-policy parameters.

To fit these non-policy parameters to data, we used Bayesian Optimization for Likelihood-Free Inference (BOLFI), a method employed for parameter estimation in models where the likelihood function is intractable or computationally expensive to evaluate [51]. This inference method has been previously applied to RL-based simulation models of human behavior [52, 53]. BOLFI requires the user to define a function quantifying the discrepancy between simulated and observed data, and employs a Gaussian process as a surrogate model to approximate how this discrepancy function varies with model parameter values.

In our study, we used BOLFI to find non-policy parameter values minimizing the following discrepancy function:

$$\text{Discrepancy} = \log\left(\left(\sum_{i=1}^{n_{ny}} \frac{|g_i - \hat{g}_i|}{g_{max}}\right) + \left(\sum_{i=1}^{n_y} \frac{|e_i - \hat{e}_i|}{e_{max}}\right) + \left(\sum_{i=1}^{n} \frac{\left|CIT_i - \hat{CIT}_i\right|}{CIT_{max}}\right) + \left(\sum_{i=1}^{n} \frac{|v_i - \hat{v}_i|}{v_{max}}\right)\right) \tag{4}$$

where the $\hat{}$ refer to model predictions, n, n_{ny}, n_y are the number of all trials, non-yielding trials, and yielding trials for one participant.

Each term in Equation 4 represents the difference between the average values of specific behavioral metrics under different kinematic conditions, as generated by various combinations of non-policy parameter values in the simulation and as observed in the human data, which our model aims to reproduce. These metrics include the gap acceptance rate (the frequency of crossing before the second vehicle in constant-speed scenarios), early crossing rate (the frequency of crossing early vs late in yielding scenarios, CIT, and average walking speed. Dividing by the maximum observed value normalizes the differences, as each metric can have a different value range.

We initialized BOLFI with uniform distributions for all the non-policy parameters, in the ranges mentioned in Section III-B4, and ran 80 optimization iterations, which was sufficient to achieve convergence for all model variants.

6) Behavioral pattern analysis: The primary goal of this analysis is to evaluate the similarity between simulated and experimental outcomes. We focused on behavioral patterns of the dependent variables—gap acceptance, early crossing rate, crossing initiation time (CIT), and average walking speed—as functions of the independent variables: time of day, time gap, vehicle speed, and eHMI. To identify statistically significant behavioral patterns in the experimental data, we employed Generalized Linear Mixed Models (GLMM) for binary dependent variables (gap acceptance and early crossing rate) and Linear Mixed Models (LMM) for continuous dependent variables (CIT and average walking speed). The fixed effects variables included *time of day (day / night)*, *Time Gap (3 s / 5 s)*, *Speed (25 mph / 30 mph)*, and *eHMI (on / off)* where applicable. Participant ID was included as a random effect variable in all models. As we observed different pedestrian behaviors in early and late crossing in yielding scenarios, we used early or late crossing as an additional fixed effect in the analysis of average crossing speed, coding 0 for late crossing and 1 for early crossing.

To evaluate whether our simulation model captures these behavioral patterns, we generate simulated datasets using the parameter-fitted model, ensuring that the sample size matches that of the original experiment. We then apply the same statistical analysis to the simulated data and compare the resulting patterns with those observed in the experimental data. This approach allows us to systematically assess the extent to which the simulated behavior aligns with real-world observations.

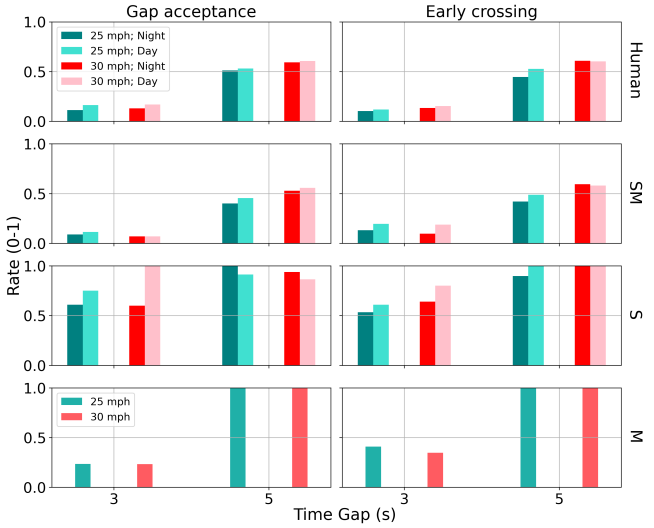


Fig. 4. Gap acceptance rate in non-yielding scenarios and early crossing rate in yielding scenarios as observed in human behavior and predicted by the models.

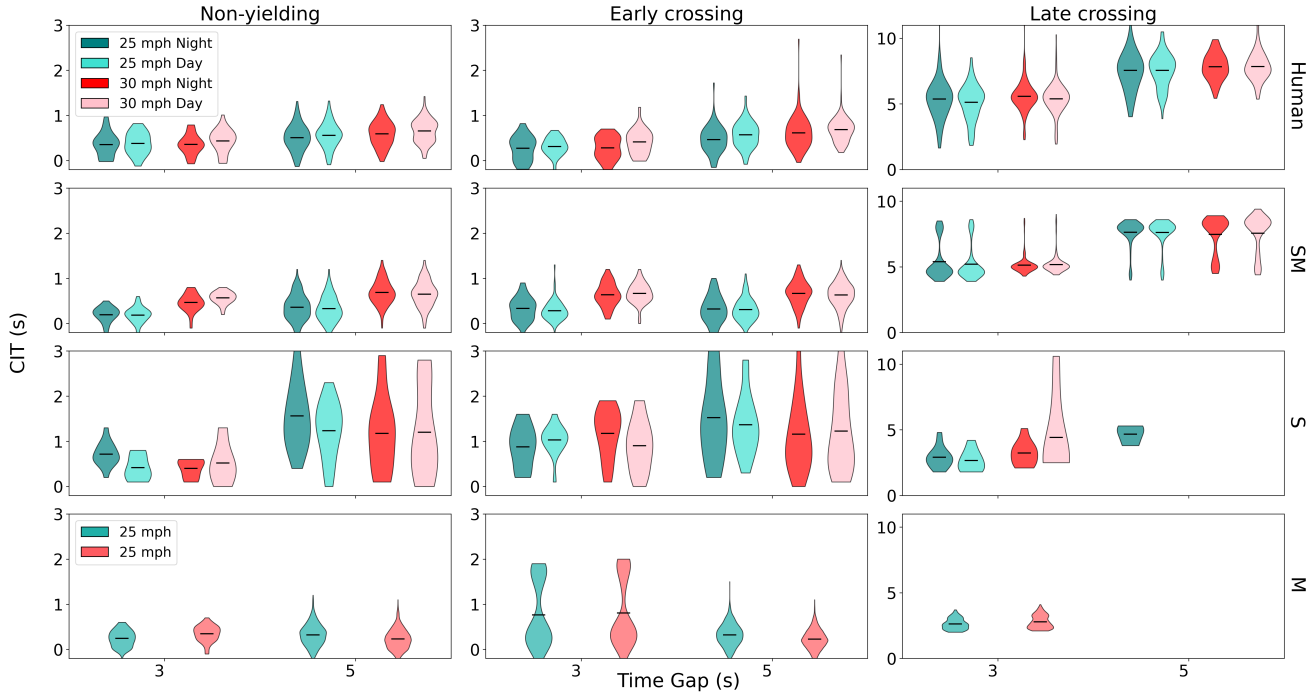


Fig. 5. Crossing Initiation Time (CIT) under different conditions. This figure shows the distribution of CIT under different conditions. The first row shows human data, the second row presents the VLW model result (accounting for walking effort), and the third row shows the VL model result (without walking effort). The blue lines within each violin plot represent the mean values. Notably, in the late crossing category for the VL model, no CIT data is available for a 5-second time gap at 30 mph, indicating that no agents chose to cross late under this condition.

TABLE III
STATISTICAL ANALYSIS OF GAP ACCEPTANCE RATE IN NON-YIELDING SCENARIOS AND EARLY CROSSING RATE IN YIELDING SCENARIOS IN THE EXPERIMENT. (* $p < 0.05$, ** $0.01 > p > 0.001$, *** $p < 0.001$).

	Gap acceptance				Early Crossing			
	Estimate	Std. Error	p-value		Estimate	Std. Error	p-value	
Intercept	-19.409	1.965	$< .001$	***	17.817	1.627	$< .001$	***
Speed	0.371	0.09926	$< .001$	***	0.426	0.090	$< .001$	***
Time Gap	3.089	0.244	$< .001$	***	2.576	0.173	$< .001$	***
Time of Day	-0.45262	0.217	0.037	*	-0.410	0.194	0.035	*
eHMI	/	/	/		-0.330	0.194	0.0885	

IV. RESULTS

All the behavioral patterns we focus on in this study are summarized in Fig. 3, with the gray cells indicating the statistically significant effects of the independent variable on pedestrian behaviors observed in the experiment. The cells with symbols mean that the corresponding model variant captures that behavioral pattern.

A. Experimental results

In this section, we present the experimental findings for the four behavioral metrics introduced in Section III-B5: gap acceptance rate in non-yielding scenarios, early crossing rate in yielding scenarios, CIT and average walking speed. We only report statistically significant effects ($p < 0.05$) in this section.

1) Gap acceptance and early crossing rate: As can be seen in the first row of Fig. 4, the gap acceptance rate and early crossing rate showed the same pattern: Pedestrians were more likely to cross when approaching vehicles were travelling at 30 mph compared to 25 mph ($p < .001$), and to cross with a larger time gap ($p < .001$). These trends are consistent with previous studies mentioned in Section II. Additionally, pedestrians were more inclined to cross during the daytime compared to nighttime ($p < 0.05$).

2) Crossing initiation time (CIT): As can be seen in the first row of Fig. 5, pedestrians were more likely to cross earlier when the speed of approaching vehicles was higher ($p < .001$), and when the time gap was greater ($p < .001$). These trends were consistent across the three categories. In addition, pedestrians showed a smaller CIT at nighttime compared to daytime in non-yielding and early crossing in yielding scenarios. Conversely, in late crossing in yielding scenarios, pedestrians initiated crossing more slowly at nighttime compared to daytime. Furthermore, from TABLE IV, eHMI had an effect on CIT in late crossing scenarios ($p < .001$). Specifically, when eHMI was on, pedestrians started to cross earlier.

3) Average walking speed: From the first row of Fig. 6, it can be seen that pedestrians crossed the road faster when the time gap was smaller ($p < .001$). Pedestrians were also more likely to cross faster during early crossings ($p < .001$). Furthermore, the average crossing speed was also affected by the eHMI ($p < .001$). Specifically, when eHMI was on, pedestrians walked slower.

Overall, the experimental data showed 19 different statistically significant effects, shown as cells shaded gray in Fig. 3.

B. Model results

In this section we examine to what extent the three different model variants were able to capture these observed effects; this is also summarized in Fig. 3. The overall collision rate in the experiment was 0.002, indicating that collisions were rare. The collision rates for the SM, S, and M models were

TABLE IV
STATISTICAL ANALYSIS OF CROSSING INITIATION TIME IN THE EXPERIMENT.

	Non-yielding				Early Crossing				Late Crossing			
	Estimate	Std. Error	p-value		Estimate	Std. Error	p-value		Estimate	Std. Error	p-value	
Intercept	-0.439	0.099	$< .001$	***	-1.514	0.402	$< .001$	***	1.435	0.425	$< .001$	***
Speed	0.042	0.006	$< .001$	***	0.099	0.028	$< .001$	***	0.094	0.030	0.001	**
Time Gap	0.105	0.009	$< .001$	***	0.197	0.041	$< .001$	***	1.006	0.037	$< .001$	***
Time of Day	-0.062	0.014	$< .001$	***	-0.151	0.061	0.014	*	0.193	0.066	0.003	*
eHMI					-0.105	0.061	0.087		-0.581	0.066	$< .001$	***

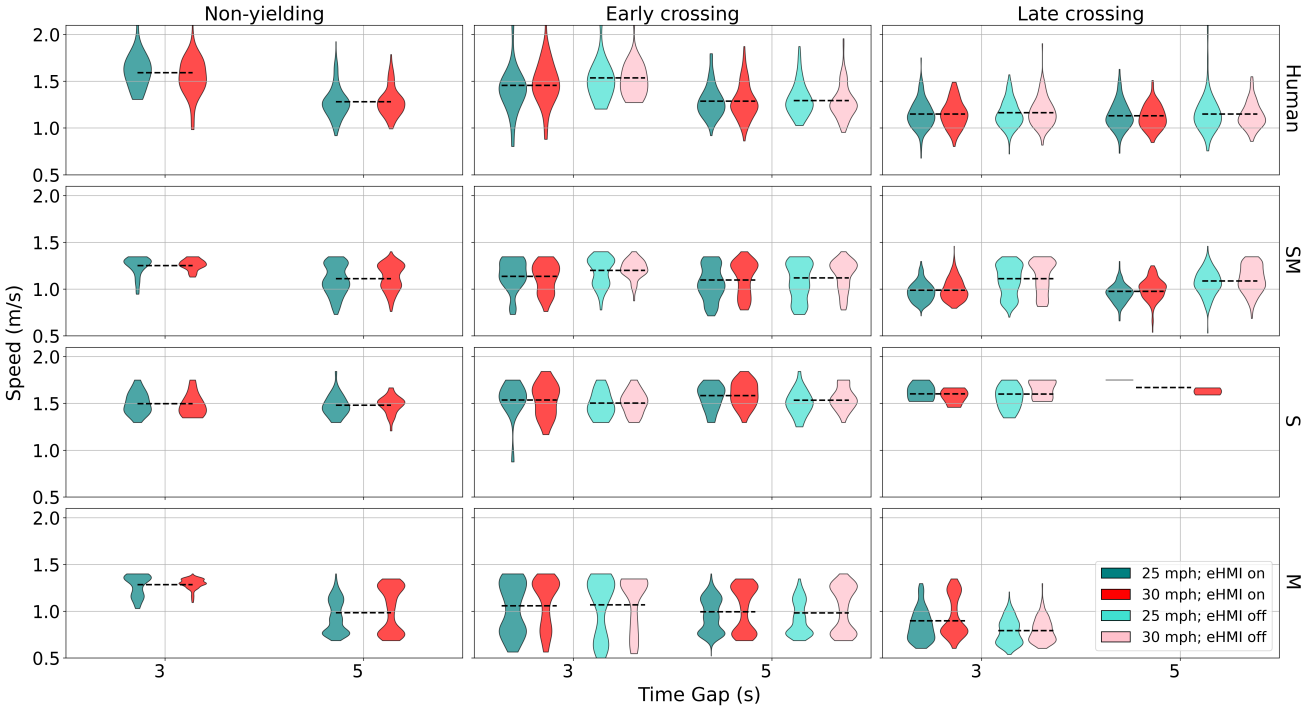


Fig. 6. Average walking speed. This figure illustrates the average walking speeds for different crossing decisions—non-yielding, early crossing, and late crossing—under different vehicle speeds (25 mph and 30 mph) and eHMI conditions (on and off). Each pair of violins represents different vehicle speeds (25 mph and 30 mph) within the same time gap and eHMI on/off condition. The black dashed lines indicate the mean walking speed for each pair of violins. This pairing method is used because the time of day did not have a significant effect on walking speed in the experiment.

TABLE V
STATISTICAL ANALYSIS OF PEDESTRIAN CROSSING SPEEDS IN THE EXPERIMENT

	Estimate	Std. Error	p-value	
Intercept	1.359	0.037	$< .001$	***
Speed	-0.001	0.002	0.609	
Time Gap	-0.046	0.003	$< .001$	***
Time of Day	-0.004	0.005	0.447	
Early cross	0.188	0.008	$< .001$	***
eHMI	-0.013	0.005	0.016	*

somewhat higher but still small, at 0.027, 0.044, and 0.040, respectively. These values indicate that each model variant could successfully learn to cross the road, with a relatively low incidence of collisions.

1) Gap acceptance and early crossing rate: Rows 2-4 of Fig. 4 show the gap acceptance behavior of the three model variants: The experimental results for gap acceptance and early crossing rates were best captured by the SM model. This model captured all six behavioral patterns. In contrast, the S and M models only captured two of these patterns: the increases with increasing time gaps, but not the speed or day/night effects. From Fig. 4, we can find that the S model shows a higher crossing tendency under short time gaps than the M model. This can be attributed to the lack of walking effort and ballistic control assumptions in the S model, enabling the agent to change speeds at each step without penalties. In contrast, the M model integrates these assumptions, which means that the agent of the M model must consider the energetic cost of accelerating or decelerating when choosing speeds. As a result, this model tends to demonstrate more conservative behavior under similar conditions, leading to a lower gap acceptance rate.

2) CIT: Rows 2-4 of Fig. 5 show the CIT of the three model variants: The SM model exhibited the best performance by accurately capturing four key behavioral patterns. It shows the increase in CIT with both speed and time gaps in non-yielding conditions, aligning closely with human data. In early crossing conditions within yielding scenarios, it captured the speed-dependent pattern, and in late crossing scenarios and illustrated the relationship between CIT and time gaps in late crossing scenarios. The S model captured two behavioral patterns: the increase in CIT with time gaps in non-yielding and early crossing conditions but failed to capture the patterns observed in late crossing scenarios or the effects of daytime and eHMI.

The M model only showed the increase in CIT with vehicle speed.

3) Average walking speed: Rows 2-4 of Fig. 6 show the average walking speed of the three model variants: The SM model captured all three behavioral patterns successfully. The M model captured two behavioral patterns: the pattern where the agent walked slower in late crossing conditions and with larger time gaps. However, the S model did not capture any of the observed patterns related to average crossing speed.

4) The pedestrian speed profile: Fig. 7 shows the speed profile of the experiment and the three model variants. It can be seen that the integration of walking effort and ballistic speed control in the SM and M models generated smooth, bell-shaped, human-like walking speed curves. These models captured the gradual acceleration and deceleration phases, resulting in a more realistic representation of pedestrian behavior. Whereas, for the S model, shown in the third row of Fig. 7, using direct speed control [46, 47, 48] rather than the biomechanical model of walking, there were abrupt fluctuations in the agent's walking speed. This is unrealistic for human locomotion, and as a result, the human-like speed profile was not captured by this model.

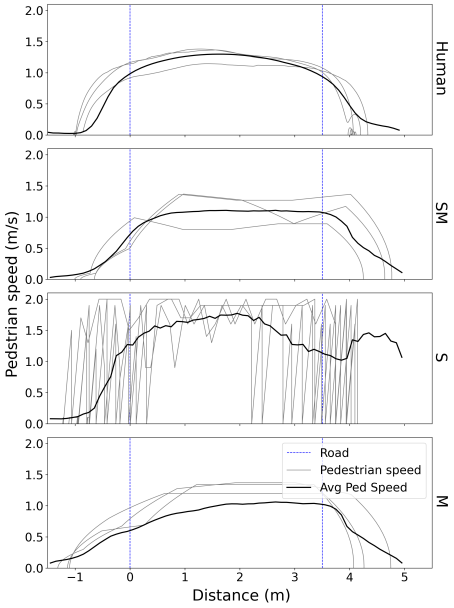


Fig. 7. Average walking speed profile of the experiment, and different model variants. Gray lines are the randomly selected speed curves and black lines are the averaged speed curves of all conditions. The blue dashed lines denote the road curbs, and the walking direction is from left to right.

V. Discussion

A. Main findings

In this study, we developed a pedestrian crossing decision model that integrates RL with sensory-motor mechanisms to simulate human-like road crossing behavior. We identified three main findings:

First, as shown in Fig. 3, the SM model variant captures most of the behavioral patterns observed in the human data, including the gap acceptance and early crossing behavior, the relation between the CIT and time gap and vehicle speed, and the behavioral patterns related to the average crossing speed.

Second, unlike previous models, considering the cognitive process, which mostly focused on the binary 'go' or 'not go' decision, our model captures the entire locomotion process of crossing the road, which can be beneficial to pedestrian behavior simulation and prediction.

Third, building on our previous crossing decision model, which primarily integrated visual limitations with RL, the model proposed here successfully captured several behavioral patterns related to motor aspects of pedestrian locomotion, the day/ night difference and the effect of eHMI, without losing the ability of the simpler model to capture sensory-related behavioral patterns. This is promising, since it shows the extensibility of the overall approach of combining mechanistic modeling of sensory-motor aspects with RL, suggesting that it may be generalizable to more complex traffic scenarios. To make the model more generalizable to more complex scenarios, a more advanced perception model is needed to sense the surrounding information including multiple vehicles and infrastructures. Additionally, we have not considered the attention mechanism in our model, which is also required for a more generalizable model.

B. Impact of sensory-motor constraints on the pedestrian crossing locomotion

1) Human-like walking speed pattern: Our SM model, with the integration of walking effort and ballistic speed control mechanisms, successfully captured the observed behavioral patterns regarding average walking speed and showed similar walking speed profile. Previous studies argued that the relation between the time gap and the crossing speed was caused by the time pressure. However, when comparing the S model and M model, we found that the model with time pressure but without walking effort and ballistic speed control assumptions could not generate that speed pattern. Therefore, we argue that the time gap-dependent walking speed pattern is the result of both the time pressure and motor aspects, modelled here in terms of walking effort and ballistic speed control.

In addition, the comparison between the S and M models further highlights the necessity of modeling human motor control to accurately simulate pedestrian locomotion. An interesting question for future work is determining the level of detail needed in the motor components of pedestrian crossing models, given that human locomotion is more complex than we have modeled and that more sophisticated human biomechanical simulations have been modelled by RL [54].

2) Day-night time difference in gap acceptance and early crossing rate: The SM model captures the day-night time difference in gap acceptance and early crossing rate by fitting each participant with individual values for the non-policy parameters of perceptual noise and looming aversion. This approach allowed us to account for variations in visual perception between day and night scenarios. We initially hypothesized that higher crossing rates during the daytime were due to smaller visual noise. However, we found no clear correlation was present between σ_{v} value and gap acceptance rate, suggesting that the gap acceptance of the model depends on a non-trivial interaction between σ_{v} and one or more of the other non-policy parameters.

3) Ablation Study on Sensory-Motor Constraints: Although we have compared the results of the S and M models, each of these two models still include two of the four main model assumptions. To further understand which assumptions have the most significant effect on the model's crossing behavior, we also conducted an ablation study involving four model variants, each excluding just one of the four components. As shown in Fig. 3, it can be observed that removing either the looming aversion or the ballistic control assumption led to the loss of eight behavioral patterns that were captured by the SM model. In contrast, removing either the walking effort or the noisy perception component resulted in the loss of six behavioral patterns. These findings suggest that looming aversion and ballistic control have a more substantial influence on crossing decisions and pedestrian behavior than the other two components.

An interesting aspect of our bounded optimality approach is its interpretability compared to data-driven neural network models. In our previous work, we found that the speed-dependent gap acceptance pattern could be attributed to the noisy perception assumption [12]. However, in the current study, we observed that even without this assumption, the model was still able to generate this behavioral pattern (although not the CIT speed-dependencies captured by the SM model). The reason for this difference may be because, in the current framework, we model the entire crossing phase, meaning the gap acceptance behavior emerges as a consequence of the entire sequential decision-making process. By contrast, our earlier work focused only on the initial go or not go decision. Additionally, the current model incorporates assumptions about the motor constraints, which were not included in the earlier study. While our method is certainly more interpretable than fully data-driven models, by providing insights into what near-optimal behavior looks like under various human constraints, it may not always be immediately obvious why the obtained behavior is near-optimal. Understanding the interaction between various human constraint assumptions in our model is non-trivial, highlighting the need to further refine our framework to enhance our understanding of pedestrian behavior dynamics, potentially leading to more robust and transparent decision-making models in the future.

4) Importance of integrated sensory-motor modeling: Neither the S model nor the M model captured as many behavioral patterns as the SM model. This suggests that modeling realistic pedestrian crossing patterns requires a comprehensive understanding of the constraints faced by humans. The SM model's success in capturing a wide range of patterns can be attributed to its incorporation of both sensory and motor mechanisms. Our comparison of models with different assumptions demonstrates that only by fully grasping and modeling these complex sensory-motor interactions can we achieve a more accurate and human-like behavioral model.

C. Implications and limitations

In this study, although vehicles exhibit predefined behavior, the scenarios include both the pedestrian interacting with the yielding and non-yielding vehicle, which means the pedestrian understands both interaction types. However, the pedestrian only interacts with the vehicles that come from one direction. Therefore, our model is most applicable to scenarios where a pedestrian crosses a single-lane road with oncoming traffic from one direction. Additionally, neither our experiment nor the RL environment indicated that the pedestrian had the priority of way, and no traffic signals were present. As such, the model is best suited for uncontrolled crosswalk settings. Our model successfully captures many observed behavioral patterns in such scenario, which suggests that it may be useful for a number of potential practical applications. It can contribute to the prediction of pedestrian behavior in real-time AV algorithms, enhancing the ability of AVs to interact safely and effectively with pedestrians [6]. Furthermore, the model can be integrated into virtual testing environments, providing a valuable tool for evaluating AV systems under various pedestrian crossing scenarios [1]. This can facilitate the development of more robust and reliable AV technologies, ultimately contributing to safer road environments. Additionally, the model has broader applications in traffic safety modeling. Accurately simulating pedestrian behavior can help the design of safer road infrastructures and traffic management systems, improving overall traffic safety and efficiency [1, 6].

However, not all of the behavioral patterns observed in the empirical study were captured by our SM model. We calculated the effect size (Cohen's d and Cohen's h) for each statistically significant variable. The effect sizes for the missing behavioral patterns were relatively small (< 0.40), suggesting that these variables do not have a substantial impact on the behavioral metrics. An exception was the effect of the time gap on CIT in early crossing scenarios in yielding conditions, where discrepancies may arise from the bimodal pattern criteria derived from experimental data not aligning well with the model's bimodal patterns. Furthermore, in the empirical study, there were statistically significant differences between young and old participants in CIT for non-yielding scenarios and in CIT for late crossings in yielding scenarios [40]. Preliminary tests with the SM model indicated that it could not easily capture these effects, possibly because we have not identified the most sensitive sensory-motor mechanism affecting these age groups differently. In this study, the pedestrian agent's movement is restricted to selecting speeds in one direction only, and trajectory similarity is not the main focus of this study. Future improvement will include giving the pedestrian agent greater freedom of action, such as heading direction, to make the model generate more human-like behavior.

Additionally, the experimental data we used was collected in a VR environment, where the pedestrian might behave differently when compared to the real world. Although the pedestrian crossing behaviors observed in our pedestrian VR simulator have been shown to be largely similar to those in real-world environments [55], to improve the applicability of our model to real-world autonomous vehicle (AV) applications and increase the generalizability of the observed behaviors, we plan to extend our model to address naturalistic datasets, where more diverse scenarios are available. For example, this study focuses on the scenario where a pedestrian interacts with

two approaching vehicles. In future work, we aim to model more complex behaviors, such as crowd-crossing dynamics, by training a multi-agent policy with an advanced perception model in an environment with more road users. This approach will allow us to capture a broader range of pedestrian behaviors and interactions, thereby improving the model's utility for real-world AV applications.

VI. Conclusion

This research presents an RL model that incorporates sensory-motor mechanisms aimed at simulating human-like pedestrian crossing behavior, which is important for the safe deployment of AVs in urban environments, and also has applications in traffic safety more broadly. The experimental findings of this study reveal that pedestrians' crossing rates and walking speeds vary in response to time gaps and vehicle speeds, and differ between day and night conditions, as well as with the presence of eHMI. Our model successfully integrated a range of sensory-motor constraints, including visual limitations, looming aversion, time pressure, walking effort, and ballistic speed control, allowing it to replicate the interaction and locomotion patterns observed in the experiment. We find that empirically observed time-gap-dependent walking speed patterns can be understood as arising from a trade-off between time pressure and walking effort, captured by our model. The ability of our model to simulate a large number of observed behavioral patterns highlights the versatility of RL in modeling complex human behaviors. It provides insights into the effect of sensory-motor mechanisms on pedestrian-vehicle interactions. Notably, the model extends the behavioral patterns captured in previous studies, demonstrating its capability to generalize across more complex pedestrian behaviors and scenarios. While the model replicates many observed behavioral patterns, it has limitations, such as capturing the influence of eHMI on CIT, and some CIT patterns in yielding scenarios. Nonetheless, by providing a more accurate representation of pedestrian behavior, this research not only contributes to the field of pedestrian behavior modeling but also has the potential to improve AV algorithms and virtual testing environments, ultimately enhancing the coexistence of AVs and pedestrians in shared spaces.

References

[1] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *IEEE transactions on intelligent transportation systems*, vol. 21, no. 3, pp. 900–918, 2019.

[2] C. Xu, C. Dong, W. Yang, Y. Hou, and H. Wang, "Platoon control and external human–machine interfaces: innovations in pedestrian–autonomous vehicle interactions," *Transportmetrica*, pp. 1–32, 2024.

[3] S. S. Man, C. Huang, Q. Ye, F. Chang, and A. H. S. Chan, "Pedestrians' interaction with ehmi-equipped autonomous vehicles: A bibliometric analysis and systematic review," *Accident Analysis & Prevention*, vol. 209, p. 107826, 2025.

[4] R. Izquierdo, J. Alonso, O. Benderius, M. Ángel Sotelo, and D. F. Llorca, "Pedestrian and passenger interaction with autonomous vehicles: Field study in a crosswalk scenario," *International Journal of Human-computer Interaction*, pp. 1–19, 2024.

[5] A. Alhawiti, V. Kwigizile, J.-S. Oh, Z. D. Asher, O. Hakimi, S. Aljohani, and S. Ayantayo, "The effectiveness of ehmi displays on pedestrian–autonomous vehicle interaction in mixed-traffic environments," *Sensors*, vol. 24, no. 15, pp. 5018–5018, 2024.

[6] F. Camara, N. Bellotto, S. Cosar, F. Weber, D. Nathanael, M. Althoff, J. Wu, J. Ruenz, A. Dietrich, G. Markkula *et al.*, "Pedestrian models for autonomous driving part ii: high-level models of human behavior," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 9, pp. 5453–5472, 2020.

[7] C. Hollmann, "A cognitive human behaviour model for pedestrian behaviour simulation," Ph.D. dissertation, University of Greenwich, 2015.

[8] G. Markkula, Y.-S. Lin, A. R. Srinivasan, J. Billington, M. Leonetti, A. H. Kalantari, Y. Yang, Y. M. Lee, R. Madigan, and N. Merat, "Explaining human interactions on the road by large-scale integration of computational psychological theory," *PNAS nexus*, vol. 2, no. 6, p. 163, 2023.

[9] V. Papathanasopoulou, I. Spyropoulou, H. Perakis, V. Gikas, and E. Andrikopoulou, "A data-driven model for pedestrian behavior classification and trajectory prediction," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 328–339, 2022.

[10] R. Quan, L. Zhu, Y. Wu, and Y. Yang, "Holistic lstm for pedestrian trajectory prediction," *IEEE transactions on image processing*, vol. 30, pp. 3229–3239, 2021.

[11] D. E. Benrachou, S. Glaser, M. Elhenawy, and A. Rakotonirainy, "Use of social interaction and intention to improve motion prediction within automated vehicle framework: A review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 22 807–22 837, 2022.

[12] Y. Wang, A. R. Srinivasan, J. P. Jokinen, A. Oulasvirta, and G. Markkula, "Pedestrian crossing decisions can be explained by bounded optimal decision-making under noisy visual perception," *Transportation Research Part C: Emerging Technologies*, vol. 171, p. 104963, 2025.

[13] K. Tian, G. Markkula, C. Wei, Y. M. Lee, R. Madigan, N. Merat, and R. Romano, "Explaining unsafe pedestrian road crossing behaviours using a psychophysics-based gap acceptance model," *Safety science*, vol. 154, p. 105837, 2022.

[14] S. Kalantarov, R. Riemer, and T. Oron-Gilad, "Pedestrians' road crossing decisions and body parts' movements," *Transportation research part F: traffic psychology and behaviour*, vol. 53, pp. 155–171, 2018.

[15] A. A. Faisal, L. P. Selen, and D. M. Wolpert, "Noise in the nervous system," *Nature reviews neuroscience*, vol. 9, no. 4, pp. 292–303, 2008.

[16] T. S. Manning, J. W. Pillow, B. Rokers, and E. A. Cooper, "Humans make non-ideal inferences about world

motion," *Journal of Vision*, vol. 22, no. 14, pp. 4054–4054, 2022.

[17] I. Kotseruba and A. Rasouli, "Intend-wait-perceive-cross: Exploring the effects of perceptual limitations on pedestrian decision-making," in *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2023, pp. 1–6.

[18] P. R. DeLucia, "Critical roles for distance, task, and motion in space perception: Initial conceptual framework and practical implications," *Human Factors*, vol. 50, no. 5, pp. 811–820, 2008.

[19] L. Mulier, H. Slabbinck, and I. Vermeir, "Face your fears: direct and indirect measurement of responses to looming threats," *Cognition and Emotion*, vol. 38, no. 1, pp. 187–197, 2024.

[20] R. E. Carlisle and A. D. Kuo, "Optimization of energy and time predicts dynamic speeds for human walking," *Elife*, vol. 12, p. e81939, 2023.

[21] T. H. Choi, "Development of a mathematical model of gait dynamics," 1997.

[22] S. Faraji, A. R. Wu, and A. J. Ijspeert, "A simple model of mechanical effects to estimate metabolic cost of human walking," *Scientific Reports*, vol. 8, no. 1, pp. 10 998–10 998, 2018.

[23] J. Oxley, B. Corben, B. Fildes, and J. Charlton, "Older pedestrians: Meeting their safety and mobility needs," in *Proceedings of the Road Safety Research, Policing and Education Conference, Perth, Australia*, 2004, pp. 14–16.

[24] A. Tustin, "The nature of the operator's response in manual control, and its implications for controller design," *Journal of the Institution of Electrical Engineers-Part IIA: Automatic Regulators and Servo Mechanisms*, vol. 94, no. 2, pp. 190–206, 1947.

[25] K. J. Craik, "Theory of the human operator in control systems. ii. man as an element in a control system," *British journal of psychology*, vol. 38, no. 3, p. 142, 1948.

[26] S. L. Barton, J. S. Matthis, and B. R. Fajen, "Control strategies for rapid, visually guided adjustments of the foot during continuous walking," *Experimental brain research*, vol. 237, pp. 1673–1690, 2019.

[27] R. Lobjois and V. Cavallo, "Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task," *Accident analysis & prevention*, vol. 39, no. 5, pp. 934–943, 2007.

[28] O. Giles, G. Markkula, J. Pekkanen, N. Yokota, N. Matsunaga, N. Merat, and T. Daimon, "At the zebra crossing: Modelling complex decision processes with variable-drift diffusion models," in *Proceedings of the 41st annual meeting of the cognitive science society*. Cognitive Science Society, 2019, pp. 366–372.

[29] J. Pekkanen, O. T. Giles, Y. M. Lee, R. Madigan, T. Daimon, N. Merat, and G. Markkula, "Variable-drift diffusion models of pedestrian road-crossing decisions," *Computational Brain & Behavior*, pp. 1–21, 2022.

[30] K. Madala and C. A. Gonzalez, "Metrics for machine learning models to facilitate sotif analysis in autonomous vehicles," *SAE International Journal of Advances and Current Practices in Mobility*, vol. 6, no. 2023-01-0829, pp. 782–790, 2023.

[31] M. Roshdi, J. Petzold, M. Wahby, H. Ebrahim, M. Berekovic, and H. Hamann, "On the road to clarity: Exploring explainable ai for world models in a driver assistance system," in *2024 IEEE Conference on Artificial Intelligence (CAI)*, 2024, pp. 1032–1039.

[32] A. R. Srinivasan, Y.-S. Lin, M. Antonello, A. Knittel, M. Hasan, M. Hawasly, J. Redford, S. Ramamoorthy, M. Leonetti, J. Billington *et al.*, "Beyond rmse: Do machine-learned models of road user interaction produce human-like behavior?" *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[33] Y. Xu, X. Yang, L. Gong, H.-C. Lin, T.-Y. Wu, Y. Li, and N. Vasconcelos, "Explainable object-induced action decision for autonomous vehicles," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 9523–9532.

[34] C. A. Diaz-Ruiz, Y. Xia, Y. You, J. Nino, J. Chen, J. Monica, X. Chen, K. Luo, Y. Wang, M. Emond *et al.*, "Ithaca365: Dataset and driving perception under repeated and challenging weather conditions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 21 383–21 392.

[35] S. Hoogendoorn and P. HL Bovy, "Simulation of pedestrian flows by optimal control and differential games," *Optimal control applications and methods*, vol. 24, no. 3, pp. 153–172, 2003.

[36] Y. Wang, A. R. Srinivasan, J. P. Jokinen, A. Oulasvirta, and G. Markkula, "Modeling human road crossing decisions as reward maximization with visual perception limitations," in *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2023, pp. 1–6.

[37] S. J. Gershman, E. J. Horvitz, and J. B. Tenenbaum, "Computational rationality: A converging paradigm for intelligence in brains, minds, and machines," *Science*, vol. 349, no. 6245, pp. 273–278, 2015.

[38] A. Oulasvirta, J. P. Jokinen, and A. Howes, "Computational rationality as a theory of interaction," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–14.

[39] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

[40] Y. M. Lee, R. Madigan, Y. Wang, H. Qin, A. Srinivasan, G. Markkula, and N. Merat, "Under the moonlight: The effect of vehicle kinematics and ehmi on older pedestrians' crossing behaviour," *PsyArXiv preprint osf.io/preprints/psyarxiv/a2m9w*, 2024.

[41] G. Markkula, J. Engström, J. Lodin, J. Bärgman, and T. Victor, "A farewell to brake reaction times? kinematics-dependent brake response in naturalistic rear-end emergencies," *Accident Analysis & Prevention*, vol. 95, pp. 209–226, 2016.

[42] P. R. DeLucia, *Perception of Collision*, ser. Cambridge Handbooks in Psychology. Cambridge University Press, 2015, p. 568–591.

[43] D. Grieve, "Gait patterns and the speed of walking." *Biomed. Eng.*, vol. 3, pp. 119–112, 1968.

[44] A. Howes, J. P. Jokinen, and A. Oulasvirta, "Towards machines that understand people," *AI Magazine*, vol. 44, no. 3, pp. 312–327, 2023.

[45] Z. Li, Y.-J. Ko, A. Putkonen, S. Feiz, V. Ashok, I. Ramakrishnan, A. Oulasvirta, and X. Bi, "Modeling touch-based menu selection performance of blind users via reinforcement learning," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: https://doi.org/10.1145/3544548.3580640

[46] G. Vizzari and T. Cecconello, "Pedestrian simulation with reinforcement learning: a curriculum-based approach," *Future Internet*, vol. 15, no. 1, p. 12, 2022.

[47] L. Sun, J. Zhai, and W. Qin, "Crowd navigation in an unknown and dynamic environment based on deep reinforcement learning," *IEEE Access*, vol. 7, pp. 109 544–109 554, 2019.

[48] D. Xu, X. Huang, Z. Li, and X. Li, "Local motion simulation using deep reinforcement learning," *Transactions in GIS*, vol. 24, no. 3, pp. 756–779, 2020.

[49] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[50] A. Raffin, A. Hill, A. Gleave, A. Kanervisto, M. Ernestus, and N. Dormann, "Stable-baselines3: Reliable reinforcement learning implementations," *Journal of Machine Learning Research*, vol. 22, no. 268, pp. 1–8, 2021. [Online]. Available: http://jmlr.org/papers/v22/20-1364.html

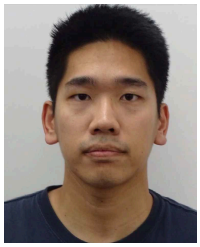
[51] M. U. Gutmann, J. Cor *et al.*, "Bayesian optimization for likelihood-free inference of simulator-based statistical models," *Journal of Machine Learning Research*, vol. 17, no. 125, pp. 1–47, 2016.

[52] A. Kangasrääsiö, K. Athukorala, A. Howes, J. Corander, S. Kaski, and A. Oulasvirta, "Inferring cognitive models from data using approximate bayesian computation," in *Proceedings of the 2017 CHI conference on human factors in computing systems*, 2017, pp. 1295–1306.

[53] H.-S. Moon, S. Do, W. Kim, J. Seo, M. Chang, and B. Lee, "Speeding up inference with user simulators through policy modulation," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–21.

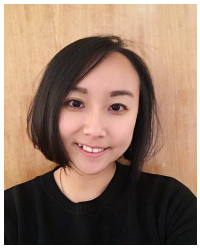
[54] S. Song, Ł. Kidziński, X. B. Peng, C. Ong, J. Hicks, S. Levine, C. G. Atkeson, and S. L. Delp, "Deep reinforcement learning for modeling human locomotion control in neuromechanical simulation," *Journal of neuroengineering and rehabilitation*, vol. 18, pp. 1–17, 2021.

[55] A. H. Kalantari, Y.-S. Lin, A. Mohammadi, N. Merat, and G. Markkula, "Testing the validity of multi participant distributed simulation for understanding and modeling road user interaction," Sep 2023. [Online]. Available: https://osf.io/preprints/psyarxiv/gk9af



Yueyang Wang received his M.Sc. in Automotive Engineering from the University of Leeds, UK, where he is currently advancing towards a Ph.D. in Transport Studies. His research intersects human factors and safety, with a strong focus on computational modeling of road user behavior and reinforcement learning.

Aravinda Ramakrishnan Srinivasan received the B.Tech. degree in electronics and communication engineering from the SASTRA University, Tirumalaisamudram, Tamil Nadu, India, and the M.S. and Ph.D. degrees in mechatronics and mechanical engineering from the University of Tennessee, Knoxville, TN, USA. He joined as the Human Factors and Traffic Safety group lead at the Institute for Highway Engineering (ISAC), RWTH Aachen University, Germany in September 2024. Before that he was research fellow at the Human Factors and Safety group, Institute for Transport Studies, University of Leeds, UK and a postdoctoral research fellow at the Lincoln Centre for Autonomous Systems, University of Lincoln, UK. His research interests include machine-learning, artificial intelligence, autonomous vehicles, and robotics applications in everyday life.



Yee Mun Lee is currently an associate professor at the Institute for Transport Studies, University of Leeds. She obtained her BSc (Hons) and her PhD degree from The University of Nottingham Malaysia. Her current research interests include investigating the interaction between automated vehicles and other road users using various methods, especially virtual reality experimental designs. She is involved in multiple EU-funded projects and is actively involved in the International Organisation for Standardisation (ISO).

Gustav Markkula received the M.Sc. degree in engineering physics and complex adaptive systems and the Ph.D. degree in machine and vehicle systems from the Chalmers University of Technology, Gothenburg, Sweden, in 2004 and 2015, respectively. He has more than a decade of research and development experience from the automotive industry. He is currently the Chair of Applied Behavior Modeling with the Institute for Transport Studies, University of Leeds, U.K. His current research interests include quantitative modeling of road user behavior and interaction, and virtual testing of vehicle safety and automation technology.